

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2003, 01:54:58 ; Search time 69 Seconds

(without alignments)
27.036 Million cell updates/sec

Handentered

Title: 70
Perfect score: 1
Sequence: 1 supprek1p1s1k 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	70	100.0	14	AA11461	Human brain carbox
2	70	100.0	246	AA11459	Human brain carbox
3	70	100.0	338	AA11458	Human brain carbox
4	70	100.0	360	AA11457	Human brain carbox
5	42	60.0	403	AA11457	Human brain carbox
6	42	60.0	431	AA11457	Human brain carbox
7	42	60.0	450	AA11457	Human brain carbox
8	42	60.0	450	AA11457	Human brain carbox
9	42	60.0	1544	AA11457	Human brain carbox
10	41	58.6	2681	AA11457	Human brain carbox

11	40	57.1	48	22	AA11461
12	40	57.1	105	21	AA11461
13	40	57.1	412	21	AA11461
14	40	57.1	904	22	AA11461
15	39	55.7	48	22	AA11461
16	39	55.7	52	22	AA11461
17	39	55.7	57	21	AA11461
18	39	55.7	58	21	AA11461
19	39	55.7	89	20	AA11461
20	39	55.7	103	22	AA11461
21	39	55.7	111	22	AA11461
22	39	55.7	230	22	AA11461
23	39	55.7	255	22	AA11461
24	39	55.7	278	22	AA11461
25	39	55.7	307	22	AA11461
26	39	55.7	317	22	AA11461
27	39	55.7	317	22	AA11461
28	39	55.7	322	21	AA11461
29	39	55.7	331	22	AA11461
30	39	55.7	442	22	AA11461
31	39	55.7	653	17	AA11461
32	39	55.7	992	22	AA11461
33	38	54.3	67	23	AA11461
34	38	54.3	69	20	AA11461
35	38	54.3	106	23	AA11461
36	38	54.3	219	23	AA11461
37	38	54.3	221	21	AA11461
38	38	54.3	282	22	AA11461
39	38	54.3	292	22	AA11461
40	38	54.3	348	21	AA11461
41	38	54.3	348	21	AA11461
42	38	54.3	359	23	AA11461
43	38	54.3	371	21	AA11461
44	38	54.3	371	21	AA11461
45	38	54.3	386	22	AA11461

ALIGNMENTS

RESULT 1	AA11461	standard; Protein: 14 AA.
ID	AA11461	
AC	AA11461	
XX	01-MAR-2001	(first entry)
DE	Human brain carboxypeptidase B protein SEQ ID NO 9.	
XX	Carboxypeptidase B; human; brain; hippocampus; APP cleavage; peptidase;	
KW	treatment; Alzheimer's disease; beta-amyloid precursor protein; vaccine;	
KW	cerebroprotective; antialzheimer's; neurotrophic; neuroprotective;	
KW	hemostatic; gene therapy; senile dementia; inherited cerebral hemorrhage;	
KW	Down's syndrome; head trauma.	
OS	Homo sapiens.	
XX	WO200066717-A1.	
XX	09-NOV-2000.	
PD	01-MAY-2000; 2000WO-JP02878.	
XX	30-APR-1999; 99JP-0125169.	
PR	(MATS/) MATSUMOTO A.	
PA	Matsumoto A;	
PI	WPI: 2000-687534/67.	
XX	Human brain carboxypeptidase B isolated from the hippocampus useful for	
PF		

PT screening agents for the treatment of Alzheimer's and other brain
PT disorders -
XX
PS Example 4; Page 78; 84pp; Japanese.
XX
CC This invention describes a novel protein with peptidase activity
CC against brain beta-amyloid precursor protein which has been isolated from
CC human hippocampus and which has cerebroprotective, antialzheimers,
CC nootropic, neuroprotective and hemostatic activity, and which can be used
CC as a vaccine or for gene therapy. The protein, and compounds identified
CC by screening as promoters or inhibitors of its activity, are used to
CC regulate beta-amyloid accumulation in the brain and treat or prevent
CC diseases in which this occurs, such as Alzheimer's, senile dementia,
CC inherited cerebral hemorrhage, Down's syndrome, and head trauma.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 70; DB 21; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNPEVKLPLSLK 14
Db 1 SNPEVKLPLSLK 14
|||||

RESULT 2
ID AAB11459 standard; Protein; 246 AA.
XX
AC AAB11459;
XX
DT 01-MAR-2001 (first entry)
XX
DE Human brain carboxypeptidase B protein SEQ ID NO 4.
XX
KW Carboxypeptidase B; human; brain; hippocampus; APP cleavage; peptidase;
KW treatment; Alzheimer's disease; beta-amyloid precursor protein; vaccine;
KW cerebroprotective; antialzheimers; nootropic; neuroprotective;
KW hemostatic; gene therapy; senile dementia; inherited cerebral hemorrhage;
KW Down's syndrome; head trauma.
XX
PI Matsumoto A;
XX
OS Homo sapiens.
XX
PN WO200066717-A1.
XX
PD 09-NOV-2000.
XX
PE 01-MAY-2000; 2000MO-JP02878.
XX
PS 30-APR-1999; 99JP-0125169.
XX
PA (MATS/) MATSUMOTO A.
XX
PI Matsumoto A;
XX
DR WPI; 2000-687534/67.
XX
PT Human brain carboxypeptidase B isolated from the hippocampus useful for
PT screening agents for the treatment of Alzheimer's and other brain
PT disorders -
XX
PS Disclosure; Page 74-75; 84pp; Japanese.
XX
CC This invention describes a novel protein with peptidase activity
CC against brain beta-amyloid precursor protein which has been isolated from
CC human hippocampus and which has cerebroprotective, antialzheimers,
CC nootropic, neuroprotective and hemostatic activity, and which can be used
CC as a vaccine or for gene therapy. The protein, and compounds identified
CC by screening as promoters or inhibitors of its activity, are used to
CC regulate beta-amyloid accumulation in the brain and treat or prevent
CC diseases in which this occurs, such as Alzheimer's, senile dementia,
CC inherited cerebral hemorrhage, Down's syndrome, and head trauma.

XX
SQ Sequence 246 AA;

Query Match 100.0%; Score 70; DB 21; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNPEVKLPLSLK 14
Db 233 SNPEVKLPLSLK 246
|||||

RESULT 3
ID AAB11458 standard; Protein; 338 AA.
XX
AC AAB11458;
XX
DT 01-MAR-2001 (first entry)
XX
DE Human brain carboxypeptidase B protein SEQ ID NO 3.
XX
KW Carboxypeptidase B; human; brain; hippocampus; APP cleavage; peptidase;
KW treatment; Alzheimer's disease; beta-amyloid precursor protein; vaccine;
KW cerebroprotective; antialzheimers; nootropic; neuroprotective;
KW hemostatic; gene therapy; senile dementia; inherited cerebral hemorrhage;
KW Down's syndrome; head trauma.
XX
PI Matsumoto A;
XX
OS Homo sapiens.
XX
PN WO200066717-A1.
XX
PD 09-NOV-2000.
XX
PE 01-MAY-2000; 2000MO-JP02878.
XX
PS 30-APR-1999; 99JP-0125169.
XX
PA (MATS/) MATSUMOTO A.
XX
PI Matsumoto A;
XX
DR WPI; 2000-687534/67.
XX
PT Human brain carboxypeptidase B isolated from the hippocampus useful for
PT screening agents for the treatment of Alzheimer's and other brain
PT disorders -
XX
PS Disclosure; Page 71-73; 84pp; Japanese.
XX
CC This invention describes a novel protein with peptidase activity
CC against brain beta-amyloid precursor protein which has been isolated from
CC human hippocampus and which has cerebroprotective, antialzheimers,
CC nootropic, neuroprotective and hemostatic activity, and which can be used
CC as a vaccine or for gene therapy. The protein, and compounds identified
CC by screening as promoters or inhibitors of its activity, are used to
CC regulate beta-amyloid accumulation in the brain and treat or prevent
CC diseases in which this occurs, such as Alzheimer's, senile dementia,
CC inherited cerebral hemorrhage, Down's syndrome, and head trauma.
XX
SQ Sequence 338 AA;

Query Match 100.0%; Score 70; DB 21; Length 338;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNPEVKLPLSLK 14
Db 325 SNPEVKLPLSLK 338
|||||

RESULT 4
ID AAB11457

ID	AA811457 standard; Protein; 360 AA.
XX	
AC	AB811457;
XX	
DT	01-MAR-2001 (first entry)
XX	
DE	Human brain carboxypeptidase B protein.
KW	Carboxypeptidase B; human; brain; hippocampus; APP cleavage; peptidase;
KM	treatment; Alzheimer's disease; beta-amyloid precursor protein; vaccine;
KW	cerebroprotective; antialzheimer's; neurotropic; neuroprotective;
KM	hemostatic; gene therapy; senile dementia; inherited cerebral hemorrhage;
KW	Down's syndrome; head trauma.
OS	Homo sapiens.
PN	MO20006717-A1.
PD	09-NOV-2000.
PE	01-MAY-2000; 2000WO-JP02878.
PR	30-APR-1999; 99JP-0125169.
PA	(MATS/) MATSUMOTO A. /
PI	Matsumoto A;
DR	WPI: 2000-687534/67.
DR	N-PSDB: MAC81962.
PT	Human brain carboxypeptidase B isolated from the hippocampus useful for
PT	screening agents for the treatment of Alzheimer's and other brain
PT	disorders -
XX	
PS	Claim 1; Page 68-71; 84pp; Japanese.
CC	This invention describes a novel protein with peptidase activity
CC	against brain beta-amyloid precursor protein which has been isolated from
CC	human hippocampus and which has cerebroprotective, antialzheimer's,
CC	neurotropic, neuroprotective and hemostatic activity, and which can be used
CC	as a vaccine or for gene therapy. The protein, and compounds identified
CC	by screening as promoters or inhibitors of its activity, are used to
CC	regulate beta-amyloid accumulation in the brain and treat or prevent
CC	diseases in which this occurs, such as Alzheimer's, senile dementia,
CC	inherited cerebral hemorrhage, Down's syndrome, and head trauma.
XX	
SQ	Sequence 360 AA:
OY	Query Match 100.0%; Score 70; DB 21; Length 360;
	Best Local Similarity 100.0%; Pred. No. 0.0015;
MATCHES	14; Conservative 0; Mismatches 0; Indels 0; Gaps 0
DB	1 SNRPVEKLLPLSLK 14 347 SNRPVEKLLPLSLK 360
RESULT 5	
AAB94027	
ID	AAB94027 standard; Protein; 403 AA.
XX	
AC	AAB94027;
XX	
DT	26-JUN-2001 (first entry)
DE	Human protein sequence SEQ ID NO:14166.
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy.
OS	Homo sapiens.
PN	EP1074617-A2.

PD	07-FEB-2001.
XX	
PF	28-JUL-2000; 2000EP-0116126.
XX	
PR	29-JUL-1999; 99JP-0248036.
PR	27-AUG-1999; 99JP-0300253.
PR	11-JAN-2000; 2000JP-0118776.
PR	02-MAY-2000; 2000JP-0183767.
PR	09-JUN-2000; 2000JP-0241899.
PA	(HELI-) HELIX RES INST.
XX	
PI	Ota T., Isogai T., Nishikawa T., Hayashi K., Saito K., Yamamoto J;
PI	Ishii S., Sugiyama T., Wakamatsu A., Negai K., Otsuki T;
DR	WPI; 2001-318749/34.
PT	Primer sets for synthesizing polynucleotides, particularly the 5602
PT	full-length cDNAs defined in the specification, and for the detection
PT	and/or diagnosis of the abnormality of the proteins encoded by the
PT	full-length cDNAs -
XX	
PS	Claim 8; SEQ ID 14166; 2537pp + CD ROM; English.
XX	
CC	The present invention describes primer sets for synthesising 5602
CC	full-length cDNAs defined in the specification. Where a primer set
CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC	to the complementary strand of a polynucleotide which comprises one of
CC	the 5602 nucleotide sequences defined in the specification, where the
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC	of an oligonucleotide comprising a sequence complementary to the
CC	complementary strand of a polynucleotide which comprises a 5'-end
CC	sequence and an oligonucleotide comprising a sequence complementary to a
CC	polynucleotide which comprises a 3'-end sequence, where the
CC	oligonucleotide comprises at least 15 nucleotides and the combination of
CC	the 5' end sequence/3' end sequence is selected from those defined in
CC	the specification. The primer sets can be used in antisense therapy and
CC	in gene therapy. The primers are useful for synthesising polynucleotides,
CC	particularly full-length cDNAs. The primers are also useful for the
CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	the full-length cDNAs. The primers allow obtaining of the full-length
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC	AAH95993 represent human amino acid sequences; and AAH13629 to AAH13632
CC	represent oligonucleotides, all of which are used in the exemplification
CC	of the present invention.
XX	
SQ	Sequence 403 AA;
Query Match .	60.0%; Score 42; DB 22; Length 403;
Best Local Similarity	77.8%; Pred. No. 79;
Matches 7; Conservative	2; Mismatches 0; Indels 0; Gaps 0
DY 3 PVEKLLPL 11	
:	
DB 86 PPLEKLPL 94	
RESULT 6	
ID AAY86203	AAY86203 standard; Protein; 431 AA.
XX	
AC AAY86203;	
DT 11-APR-2000	(first entry)
XX	
DE Nuclear transport protein clone hPB2060 protein sequence.	
KW Nuclear transport protein; drug delivery system; visual detection; human;	
KW nuclear transport indicator.	
OS Homo sapiens.	

XX PN W09964455-A1.
 XX PD 16-DEC-1999.
 XX PE 04-JUN-1999; 99WO-JP03015.
 XX PR 05-JUN-1998; 98JP-0174065.
 XX PR 14-APR-1999; 99JP-0107262.
 XX PA (HELI-) HELIX RES INST.
 XX P1 Ueki N, Yano K;
 XX WPI; 2000-105872/09.
 XX DR N-PSDB; AA296761.
 XX PT Peptides with nuclear transport activity, applicable as carrier for
 PT transport desirable substances, e.g. drugs into nucleus, and as nuclear
 PT transport indicators after modification for visual detection
 XX PS Claim 1; Page 84-87; 309pp; Japanese.
 XX CC This sequence represents the amino acid sequence of a nuclear transport
 CC protein of the invention. The invention relates to peptides with nuclear
 CC transport activity, and also includes similar peptides with the same
 CC transport activity, but with some amino acids deleted, substituted and/or
 CC added. The peptides can be used as carriers for transferring desirable
 CC substances particularly drugs, including nucleic acids and proteins, into
 CC the nucleus in a drug delivery system. The peptides can also be used as
 CC nuclear transport indicators after modification for visual detection.
 CC Drugs can be delivered efficiently with the peptides to a specific target
 CC site. This delivery allows the maintenance of drug concentration,
 CC therapeutic efficacy and a reduction of side-effects of such drugs to be
 CC achieved.
 XX SQ Sequence 431 AA;
 Query Match 60.0%; Score 42; DB 21; Length 431;
 Best Local Similarity 77.8%; Pred. No. 86;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PPEVKLLPL 11
 DB 147 PPEVKLLPL 155
 RESULT 7
 ABB91564
 ID ABB91564 standard; Protein: 450 AA.
 XX AC ABB91564;
 XX DT 31-MAY-2002 (first entry)
 XX DE Herbicidally active polypeptide SEQ ID NO 775.
 XX KW Herbicidally plant; agriculture; herbicide.
 XX OS Arabidopsis thaliana.
 XX PN W0200210210-A2.
 XX PD 07-FEB-2002.
 XX PF 28-AUG-2001; 2001WO-EP09892.
 XX PR 28-AUG-2001; 2001WO-EP09892.
 XX PA (FARB) BAYER AG.
 XX PI Tietjen K, Weidler M;
 XX CC

DR WPI; 2002-269010/31.
 XX PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX PS Claim 5; SEQ ID NO 775; 261pp + Sequence Listing; English.
 XX CC The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 XX SQ Sequence 450 AA;
 Query Match 60.0%; Score 42; DB 23; Length 450;
 Best Local Similarity 75.0%; Pred. No. 90;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SNPEVKLLPL 12
 DB 26 TNPPEVKLLPL 37
 RESULT 8
 AAB94063
 ID AAB94063 standard; Protein: 1038 AA.
 XX AC AAB94063;
 XX DT 26-JUN-2001 (first entry)
 XX DE Human protein sequence SEQ ID NO:14242.
 XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX OS Homo sapiens.
 XX PN EP1074617-A2.
 XX PD 07-FEB-2001.
 XX FE 28-JUL-2000; 2000EP-0116126.
 XX PR 29-JUL-1999; 99JP-0248036.
 XX PR 27-AUG-1999; 99JP-0300253.
 XX PR 11-JAN-2000; 2000JP-0118776.
 XX PR 02-MAY-2000; 2000JP-0183767.
 XX PR 09-JUN-2000; 2000JP-0241899.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX DR WPI; 2001-318749/34.
 XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX PS Claim 8; SEQ ID 14242; 2537pp + CD ROM; English.
 XX CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
SQ Sequence 1038 AA;

Query Match 60.0%; Score 42; DB 22; Length 1038;
Best Local Similarity 77.8%; Pred. No. 2.3e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPEVKLLPL 11
II:III:III
DB 832 PPLEKILPL 840

RESULT 9
AAV41109
ID AAV41109 standard; Protein; 1544 AA.

XX AAV41109;

DT 17-JAN-2000 (first entry)

DE Human cancer-associated polypeptide plu-1.

KM Human; cancer-associated polypeptide; plu-1; ovarian cancer; vaccine;
KM breast cancer; immune response; cytotoxic T lymphocyte; imaging agent;
KM therapeutic; plu-1 antigen; breast tumor; tumor antigen.

OS Homo sapiens.

PN WO9949034-A1.

PD 30-SEP-1999.

PF 19-MAR-1999; 99WO-GB00866.

PR 20-MAR-1998; 98GB-0005877.

PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.

PI Taylor-papadimitriou J;

DR WPI: 1999-591090/50.

DR N-PSDB; AA23079.

XX New nucleic acid encoding the cancer-associated polypeptide plu-1, for
XX diagnosis, treatment and prevention of cancer, especially of breast and
XX ovary

PS Claim 11; Fig 2; 173pp; English.

XX This represents a human cancer-associated polypeptide plu-1. The plu-1
CC polypeptide can be recombinantly expressed by standard recombinant
CC methodology. Detection of the plu-1 nucleic acid or the polypeptide is
CC used for the following: (i) diagnosis (including imaging) and prognosis
CC of, and determination of susceptibility to, cancer, specifically ovarian

CC or breast cancer; and (i) treating cancer (by inducing an immune response
CC against cancer cells, e.g. as a vaccine, or by antisense inhibition).
CC Antigens derived from the polypeptide are used to generate activated
CC cytotoxic T lymphocytes, or dendritic cells, for subsequent return to the
CC patient for treatment of cancer. The polypeptide may also be used to
CC identify inhibitors of plu-1 activity. Fragments of the polypeptide, and
CC antibodies raised against plu-1, are useful as assay and imaging agents,
CC also therapeutically (to induce an anti-idiotypic response or where
CC conjugated to cytotoxic agents). The plu-1 antigen is expressed more
CC commonly in breast tumors than some known tumor antigens.
SQ Sequence 1544 AA;

Query Match 60.0%; Score 42; DB 20; Length 1544;
Best Local Similarity 77.8%; Pred. No. 3.5e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPEVKLLPL 11
II:III:III
DB 1227 PPLEKILPL 1235

RESULT 10
ABG29184
ID ABG29184 standard; Protein; 2681 AA.

XX ABG29184;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #29175.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YF;

DR WPI: 2001-639362/73.

DR N-PSDB; AAS93371.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity

PS Claim 20; SEQ ID No 59543; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (II) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in

PR	22-AUG-2000	2000US-0226868
PR	22-AUG-2000	2000US-0227189
PR	23-AUG-2000	2000US-0227000
PR	30-AUG-2000	2000US-0228984
PR	01-SEP-2000	2000US-0228287
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229509
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230477
PR	06-SEP-2000	2000US-0230478
PR	08-SEP-2000	2000US-0231242
PR	08-SEP-2000	2000US-0231243
PR	08-SEP-2000	2000US-0231244
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0231415
PR	08-SEP-2000	2000US-0232080
PR	08-SEP-2000	2000US-0232081
PR	12-SEP-2000	2000US-0231968
PR	14-SEP-2000	2000US-0232397
PR	14-SEP-2000	2000US-0232398
PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232400
PR	14-SEP-2000	2000US-0232401
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234224
PR	25-SEP-2000	2000US-0234998
PR	25-SEP-2000	2000US-0234998
PR	26-SEP-2000	2000US-0235484
PR	27-SEP-2000	2000US-0235834
PR	27-SEP-2000	2000US-0235836
PR	29-SEP-2000	2000US-0236367
PR	29-SEP-2000	2000US-0236368
PR	29-SEP-2000	2000US-0236369
PR	29-SEP-2000	2000US-0236370
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241121
PR	20-OCT-2000	2000US-0241122
PR	20-OCT-2000	2000US-0241185
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241787
PR	20-OCT-2000	2000US-0241808
PR	20-OCT-2000	2000US-0241809
PR	08-NOV-2000	2000US-0246477
PR	08-NOV-2000	2000US-0246478
PR	08-NOV-2000	2000US-0246523
PR	08-NOV-2000	2000US-0246524
PR	08-NOV-2000	2000US-0246525
PR	08-NOV-2000	2000US-0246526
PR	08-NOV-2000	2000US-0246527
PR	08-NOV-2000	2000US-0246528
PR	08-NOV-2000	2000US-0246532
PR	08-NOV-2000	2000US-0246539
PR	08-NOV-2000	2000US-0246611
PR	08-NOV-2000	2000US-0246611
PR	08-NOV-2000	2000US-0246613
PR	08-NOV-2000	2000US-0246617

PR 17-NOV-2000; 2000US-0249308.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HDMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-483426/52.
DR N-PSDB; AAK53341.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX
PS Claim 11: SEQ ID NO 10153; 3071bp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK67694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
XX
SQ Sequence 48 AA;
Query Match 57 1%; Score 40; DB 22; Length 48;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 12
AAB40486
ID AAB40486 standard; Protein; 105 AA.
XX
AC AAB40486;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF250 polypeptide sequence SEQ ID NO:500.
XX
KW Human: open reading frame; ORFX: detection; cytostatic; hepatotropic;
KW vulnerrary; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antihypoid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disease; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
XX
XX Thrombosis; contraceptive.
XX
XX Homo sapiens.
XX
XX MO200058473-A2.
XX
XX
PD 05-OCT-2000.
XX
XX
XX 31-MAR-2000; 2000MO-US08621.
XX
XX
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Leach M;
XX
XX WPI: 2000-602362/57.
DR N-PSDB; AAC74695.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX
PS Claim 11; Page 677-678; 5507bp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerrary;
CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
CC osteopathic; anticonvulsant; antirheumatic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antihypoid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 105 AA;
SQ
Query Match 57.1%; Score 40; DB 21; Length 105;
Best Local Similarity 80.0%; Pred. NO. 39;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SNPEVKLPLP 10
|||||
DB 40 SNPEVKLPLP 49

RESULT 13
AA97823
ID AA97823 standard; Protein; 412 AA.

AC AA97823;
XX 29-AUG-2000 (first entry)
XX
XX Pseudomonas sp. WF505 ORF6-6 protein, SEQ ID NO:11.
DE
XX Benzochlorophene oxidase; bco gene cluster; desulphurisation; crude oil;
KW decontamination.
XX
OS Pseudomonas sp. WF505.
XX
XX JP2000093180-A.
XX
XX 04-APR-2000.
XX
XX 28-SEP-1998; 98JP-0272744.
XX
XX 28-SEP-1998; 98JP-0272744.
XX
XX (TOFU) TONEN CORP.
XX
XX WPI: 2000-378972/33.
XX
XX N-PSDB: AAA38394.

XX
XX New proteins useful for desulfurization of crude oil have ability to
XX control expression of benzochlorophene oxidase group -
XX
XX
XX Example 1; Page 24-25; 43pp; Japanese.

XX The invention relates to proteins encoded by the benzochlorophene oxidase
XX (bco) gene cluster from Pseudomonas sp. WF505, specifically the protein
XX which controls expression of the gene cluster (AA97823), the
XX benzochlorophene oxidase (bco) enzyme itself (AA97825) and a
XX benzochlorophene oxidase co-enzyme (AA97824), which controls
XX benzochlorophene activity via redox regulation. These enzymes are
XX respectively encoded by the bco gene cluster open reading frames ORF2-16
XX (AAA38398), ORF4-9 (AAA38396) and ORF6-7 (AAA38395). The invention also
XX encompasses genetic constructs and vectors comprising the genes encoding
XX these proteins, hosts containing such constructs and the recombinant
XX production of the enzymes. Pseudomonas sp. WF505 benzochlorophene oxidase,
XX benzochlorophene oxidase co-enzyme and bco gene cluster regulator are
XX useful in the desulfurization of crude oil and petroleum products, and
XX for the clarification of soil and waste water contaminated by organic
XX sulphur compounds. Sequences AA97819-197829 represent the proteins
XX encoded by the genes of the Pseudomonas sp. WF505 benzochlorophene oxidase
XX gene cluster.
XX
XX
XX

SQ Sequence 412 AA;

Query Match 57.1%; Score 40; DB 21; Length 412;
Best Local Similarity 58.3%; Pred. NO. 1.8e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNPEVKLPLP 12
|||||
DB 297 SNPEVKLPLP 308

RESULT 14
ABB80607
ID ABB80607 standard; Protein; 904 AA.

XX ABB80607;
XX
XX 08-OCT-2002 (first entry)
XX
XX Human sbg1016951GBrecept protein #1.
XX
XX

XX Human; secreted protein; immunosuppressive; cytostatic; nocotropic;
KW neuroprotective; antitumor; vulnery; antileukemic; ophthalmological;
KW antiparkinsonian; antileukemic; antileukemic; dermatological;
KW hypotensive; cerebroprotective; virucide; antiinflammatory; diabetes;
KW malignant tumor; hypertension; hypotension; obesity; bulimia; anorexia;
KW Huntington's disease; Tourette's syndrome; delirium; mental retardation;
KW mental disorder; sexual development disorder; blood cascade dysfunction;
KW stroke; growth disorder.
KW
XX
XX

OS Homo sapiens.
XX
XX WO200222802-A1.
XX
XX 21-MAR-2002.
XX
XX
XX 13-SEP-2001; 2001WO-US28462.
XX
XX
XX 13-SEP-2000; 2000US-232455P.
XX
XX 13-SEP-2000; 2000US-232463P.
XX
XX 02-OCT-2000; 2000US-237293P.
XX
XX 07-NOV-2000; 2000US-246269P.
XX
XX 20-NOV-2000; 2000US-252049P.
XX
XX

XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX (GLAX) GLAXO GROUP LTD.
XX
XX

XX Agarwal P, Gogswell JP, Lai Y, Martensen SA, Smith RF, Strum JC;
XX Xie Q;
XX
XX

XX WPI: 2002-393963/42.
XX
XX N-PSDB: ABL57743.
XX
XX

XX Novel isolated secreted polypeptides and polynucleotides encoding them
XX useful for treating cancer, Alzheimer's disease, tumor metastasis,
XX autosomal recessive atypical hemolytic uremic syndrome, wound healing
XX disorder -
XX
XX 1; Page235-237; 246pp; English.

XX The invention relates to a novel isolated polypeptide (ABB80569-ABB80612)
XX (secreted polypeptide) which is encoded by any one of 44 polynucleotide
XX sequences (ABL57705-ABL57748) given in the specification. The
XX polypeptides have immunosuppressive, cytostatic, nocotropic,
XX neuroprotective, antitumor, vulnery, antileukemic, ophthalmological,
XX antiparkinsonian, antirheumatic, antileukemic, dermatological,
XX hypotensive, cerebroprotective, virucide, and antiinflammatory activity.
XX The polynucleotide and polypeptide are useful for treating diabetes,
XX malignant tumours, hyper- and hypotension, obesity, bulimia, anorexia,
XX asthma, manic depression, dementia, delirium, mental retardation,
XX Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental
XX or sexual development disorders, and dysfunctions of the blood cascade
XX system including those leading to stroke.
XX
XX
XX

SQ Sequence 904 AA;

Query Match 57.1%; Score 40; DB 23; Length 904;
Best Local Similarity 75.0%; Pred. NO. 4.2e+02;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SNPEVKLPLP 12

DB	435	SEPPVEDLPPKS	446
RESULT	15		
AAU86840			
ID	AAU86840	standard; Protein; 48 AA.	
XX	AAU86840;		
AC			
XX			
DT	21-MAY-2002	(first entry)	
XX			
DE	Novel human connective tissue related polypeptide #406.		
XX			
KW	Human; connective tissue related disorder; cancer; cytoslatic.		
XX			
OS	Homo sapiens.		
XX	WO200155343-A1.		
XX			
PD	02-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US01322.		
XX			
PR	31-JAN-2000; 2000US-0179065.		
PR	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184664.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0216880.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		
PR	14-AUG-2000; 2000US-0224519.		
PR	14-AUG-2000; 2000US-0225213.		
PR	14-AUG-2000; 2000US-0225214.		
PR	14-AUG-2000; 2000US-0225266.		
PR	14-AUG-2000; 2000US-0225267.		
PR	14-AUG-2000; 2000US-0225268.		
PR	14-AUG-2000; 2000US-0225270.		
PR	14-AUG-2000; 2000US-0225447.		
PR	14-AUG-2000; 2000US-0225757.		
PR	14-AUG-2000; 2000US-0225758.		
PR	14-AUG-2000; 2000US-0225759.		
PR	18-AUG-2000; 2000US-0225729.		
PR	22-AUG-2000; 2000US-0226881.		
PR	22-AUG-2000; 2000US-0226882.		
PR	23-AUG-2000; 2000US-0227182.		
PR	30-AUG-2000; 2000US-0227009.		
PR	30-AUG-2000; 2000US-0228924.		
PR	01-SEP-2000; 2000US-0228287.		
PR	01-SEP-2000; 2000US-0228343.		
PR	01-SEP-2000; 2000US-0229344.		
PR	01-SEP-2000; 2000US-0229345.		
PR	05-SEP-2000; 2000US-0229509.		
PR	05-SEP-2000; 2000US-0229513.		
PR	06-SEP-2000; 2000US-0230437.		
PR	06-SEP-2000; 2000US-0230438.		
PR	08-SEP-2000; 2000US-0231242.		
PR	08-SEP-2000; 2000US-0231243.		
PR	08-SEP-2000; 2000US-0231244.		
PR	08-SEP-2000; 2000US-0231413.		
PR	08-SEP-2000; 2000US-0231414.		
PR	08-SEP-2000; 2000US-0232080.		
PR	08-SEP-2000; 2000US-0232081.		
PR	12-SEP-2000; 2000US-0231968.		
PR	14-SEP-2000; 2000US-0232397.		
PR	14-SEP-2000; 2000US-0232398.		
PR	14-SEP-2000; 2000US-0232399.		
PR	14-SEP-2000; 2000US-0232400.		
PR	14-SEP-2000; 2000US-0232401.		
PR	14-SEP-2000; 2000US-0233063.		
PR	14-SEP-2000; 2000US-0233064.		
PR	14-SEP-2000; 2000US-0233065.		
PR	21-SEP-2000; 2000US-0234223.		
PR	21-SEP-2000; 2000US-0234274.		
PR	25-SEP-2000; 2000US-0234997.		
PR	25-SEP-2000; 2000US-0234998.		
PR	26-SEP-2000; 2000US-0235484.		
PR	27-SEP-2000; 2000US-0235834.		
PR	27-SEP-2000; 2000US-0235836.		
PR	29-SEP-2000; 2000US-0236327.		
PR	29-SEP-2000; 2000US-0236367.		
PR	29-SEP-2000; 2000US-0236368.		
PR	29-SEP-2000; 2000US-0236369.		
PR	29-SEP-2000; 2000US-0236370.		
PR	02-OCT-2000; 2000US-0236802.		
PR	02-OCT-2000; 2000US-0237037.		
PR	02-OCT-2000; 2000US-0237038.		
PR	02-OCT-2000; 2000US-0237039.		
PR	13-OCT-2000; 2000US-0237040.		
PR	13-OCT-2000; 2000US-0239935.		
PR	13-OCT-2000; 2000US-0239937.		
PR	20-OCT-2000; 2000US-0240960.		
PR	20-OCT-2000; 2000US-0241221.		
PR	20-OCT-2000; 2000US-0241785.		
PR	20-OCT-2000; 2000US-0241786.		
PR	20-OCT-2000; 2000US-0241787.		
PR	20-OCT-2000; 2000US-0241808.		
PR	20-OCT-2000; 2000US-0241809.		
PR	20-OCT-2000; 2000US-0241826.		
PR	01-NOV-2000; 2000US-0244617.		
PR	08-NOV-2000; 2000US-0246474.		
PR	08-NOV-2000; 2000US-0246475.		
PR	08-NOV-2000; 2000US-0246476.		
PR	08-NOV-2000; 2000US-0246477.		
PR	08-NOV-2000; 2000US-0246478.		
PR	08-NOV-2000; 2000US-0246523.		
PR	08-NOV-2000; 2000US-0246524.		
PR	08-NOV-2000; 2000US-0246525.		
PR	08-NOV-2000; 2000US-0246526.		
PR	08-NOV-2000; 2000US-0246527.		
PR	08-NOV-2000; 2000US-0246528.		
PR	08-NOV-2000; 2000US-0246532.		
PR	08-NOV-2000; 2000US-0246609.		
PR	08-NOV-2000; 2000US-0246610.		
PR	08-NOV-2000; 2000US-0246611.		
PR	17-NOV-2000; 2000US-0246613.		
PR	17-NOV-2000; 2000US-0246614.		
PR	17-NOV-2000; 2000US-0246615.		
PR	17-NOV-2000; 2000US-0246616.		
PR	17-NOV-2000; 2000US-0246617.		
PR	17-NOV-2000; 2000US-0249212.		
PR	17-NOV-2000; 2000US-0249213.		
PR	17-NOV-2000; 2000US-0249214.		
PR	17-NOV-2000; 2000US-0249215.		
PR	17-NOV-2000; 2000US-0249216.		
PR	17-NOV-2000; 2000US-0249217.		
PR	17-NOV-2000; 2000US-0249218.		
PR	17-NOV-2000; 2000US-0249244.		
PR	17-NOV-2000; 2000US-0249245.		
PR	17-NOV-2000; 2000US-0249264.		
PR	17-NOV-2000; 2000US-0249265.		
PR	17-NOV-2000; 2000US-0249267.		
PR	17-NOV-2000; 2000US-0249299.		

PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-565190/63.
DR N-PSDB; ABK42018.
XX
XX Nucleic acid encoding novel connective tissue associated polypeptides,
PT used in diagnosing, preventing, treating or ameliorating a disorder
PT such as cancer or rheumatoid arthritis -
XX
XX
PS Claim 11; SEQ ID No 905; 673pp; English.
XX
XX The present invention relates to the isolation of novel human
CC connective tissue related polypeptides and the polynucleotide
CC (CDNA and genomic) sequences encoding them. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of diseases associated with connective tissue(s), including
CC cancer. The polynucleotide sequences of the invention are also useful
CC in gene therapy. AU86435-AU86923 represent the novel human connective
CC tissue related polypeptides.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 48 AA;
Query Match 55.7%; Score 39; DB 22; Length 48;
Best Local Similarity 80.0%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 4 PVEKLPLSL 13
|:| |
DB 33 PLETLPLSL 42

Search completed: January 12, 2003, 04:15:03
Job time : 75 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2003, 04:14:30 ; Search time 23 Seconds

(without alignments)
17,910 Million cell updates/sec

Title: HANDENTERED
Perfect score: 70
Sequence: 1 snpveklplslk 14

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents,AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	55.7	653	1	US-08-339-152A-16
2	39	55.7	653	2	US-08-007-999B-3
3	39	55.7	653	2	US-08-689-276A-3
4	37	52.9	305	1	US-08-680-726A-80
5	37	52.9	305	4	US-09-092-409-80
6	37	52.9	724	2	US-09-562-737-25
7	36.5	52.1	687	2	US-08-540-804-4
8	36.5	52.1	687	2	US-08-218-265-4
9	36.5	52.1	687	2	US-08-521-872-4
10	36.5	52.1	687	4	US-08-590-399-4
11	36	51.4	315	4	US-08-936-165A-368
12	36	51.4	315	4	US-09-362-318-2
13	36	51.4	464	1	US-07-688-352C-16
14	36	51.4	464	2	US-08-474-379C-16
15	36	51.4	464	2	US-09-146-249A-16
16	36	51.4	464	3	US-08-206-188B-16
17	36	51.4	464	5	PCR-US91-02714-16
18	36	51.4	528	2	US-08-466-589-2
19	36	51.4	528	2	US-08-700-636-2
20	36	51.4	528	3	US-08-467-574-2
21	36	51.4	528	4	US-09-217-345-2
22	36	51.4	529	1	US-08-496-855A-2
23	36	51.4	529	4	US-08-487-596-2
24	36	51.4	600	4	US-09-423-468A-13
25	35	50.0	17	2	US-08-964-725-15
26	35	50.0	93	2	US-08-964-725-14
27	35	50.0	388	2	US-08-282-197C-56

28	35	50.0	1263	4	US-09-446-504-6	Sequence 6, Appl
29	35	50.0	1263	4	US-09-712-266-6	Sequence 6, Appl
30	35	50.0	1263	4	US-09-091-889A-4	Sequence 4, Appl
31	35	50.0	1727	2	US-08-477-451-10	Sequence 10, Appl
32	34	48.6	237	4	US-08-999-774A-8	Sequence 8, Appl
33	34	48.6	258	4	US-08-914-999-16	Sequence 16, Appl
34	34	48.6	266	4	US-08-904-234-3	Sequence 3, Appl
35	34	48.6	429	2	US-08-677-049-5	Sequence 5, Appl
36	34	48.6	463	3	US-09-082-310-1	Sequence 1, Appl
37	34	48.6	463	3	US-09-575-205-1	Sequence 5071, Ap
38	34	48.6	507	4	US-09-134-001C-5071	Sequence 32, Appl
39	34	48.6	532	1	US-08-339-152A-32	Sequence 2, Appl
40	34	48.6	548	3	US-08-942-001-2	Sequence 2, Appl
41	34	48.6	548	4	US-09-337-386-2	Sequence 2, Appl
42	34	48.6	548	4	US-09-846-922-2	Sequence 1, Appl
43	34	48.6	575	4	US-08-924-345-3	Sequence 3, Appl
44	34	48.6	640	4	US-09-177-165A-30	Sequence 30, Appl
45	34	48.6	807	4	US-09-177-650-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-08-339-152A-16
Sequence 16, Application US/08339152A
Patent No. 5643726

GENERAL INFORMATION:

APPLICANT: Tanzl, Rudolph E.

TITLE OF INVENTION: Methods For Modulating Transcription

TITLE OF INVENTION: From The Amyloid -Protein Precursor (APP) Promoter

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSER: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 New York Ave., NW, Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/339,152A

FILING DATE: 10-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Steffe, Eric K.

REGISTRATION NUMBER: 36,688

REFERENCE/DOCKET NUMBER: 0609.4120000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

TELEX:

INFORMATION FOR SEQ. ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 653 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-339-152A-16

Query Match 55.7%; Score 39; DB 1; Length 653;
Best local Similarity 72.7%; Pred. No. 77;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PVEYKLPPLSL 13
Db 17 PPLPPLPLSL 27

RESULT 2
US-08-007-999B-3
; Sequence 3, Application US/08007999B
; Patent No. 5851787
; GENERAL INFORMATION:
; APPLICANT: Masco, Wilma
; APPLICANT: Bupp, Keith
; APPLICANT: Magendantz, Margaret
; APPLICANT: Tanzl, Rudolph
; APPLICANT: Solomon, Frank
; TITLE OF INVENTION: AMYLOID PRECURSOR-LIKE PROTEIN AND USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
; STREET: 1100 New York Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/007,999B
; FILING DATE: 21-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,642
; FILING DATE: 20-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,022
; FILING DATE: 17-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, G. Kevin
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: 0609.3520002/JAG/GKT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2571
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 653 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-007-999B-3
Query Match 55.7%; Score 39; DB 2; Length 653;
Best Local Similarity 72.7%; Pred. No. 77;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 PVEKLLPLSL 13
||:|||||
Db 17 PPLPLPLSL 27

RESULT 3
US-08-689-276A-3
; Sequence 3, Application US/08689276A
; Patent No. 5891991
; GENERAL INFORMATION:
; APPLICANT: Masco, Wilma
; APPLICANT: Bupp, Keith
; APPLICANT: Magendantz, Margaret
; APPLICANT: Tanzl, Rudolph
; APPLICANT: Solomon, Frank
; TITLE OF INVENTION: AMYLOID PRECURSOR-LIKE PROTEIN AND USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/689,276A
; FILING DATE: 06-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/007,999
; FILING DATE: 21-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,642
; FILING DATE: 20-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,022
; FILING DATE: 17-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: JORGE A. GOLDSTEIN
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 0609.3520003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2571
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 653 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-689-276A-3
Query Match 55.7%; Score 39; DB 2; Length 653;
Best Local Similarity 72.7%; Pred. No. 77;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 PVEKLLPLSL 13
||:|||||
Db 17 PPLPLPLSL 27

RESULT 4
US-08-680-726A-80
; Sequence 80, Application US/08680726A
; Patent No. 5804197
; GENERAL INFORMATION:
; APPLICANT: Haanes, Elizabeth J.
; APPLICANT: Frank, Rexann S.
; TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUS
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,726A
; FILING DATE: 12-JUL-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-46-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 305 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-680-726A-80

Query Match
Best Local Similarity 52.9%; Score 37; DB 1; Length 305;
Best Local Similarity 61.5%; Pred. No. 72;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NPVEKLLPLSLK 14
: | | | | | | | | | |
Db 107 HPTVEKLNIPYDLK 119

RESULT 5
US-09-092-409-80
Sequence 80, Application US/09092409
Patent No. 6159478
GENERAL INFORMATION:
APPLICANT: Haanes, Elizabeth J.
APPLICANT: Frank, Rexann S.
TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/092,409
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/680,726
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-46-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 305 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-092-409-80

Query Match
Best Local Similarity 52.9%; Score 37; DB 4; Length 305;
Best Local Similarity 61.5%; Pred. No. 72;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NPVEKLLPLSLK 14
: | | | | | | | | | |
Db 107 HPTVEKLNIPYDLK 119

RESULT 6
US-09-562-737-25
Sequence 25, Application US/09562737
Patent No. 6428967
GENERAL INFORMATION:
APPLICANT: Herz, Joachim
APPLICANT: Gotthardt, Michael
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTS0708
CURRENT APPLICATION NUMBER: US/09/562,737
CURRENT FILING DATE: 2000-05-01
NUMBER OF SEQ ID NOS: 132
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 25
LENGTH: 724
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-25

Query Match
Best Local Similarity 52.9%; Score 37; DB 4; Length 724;
Best Local Similarity 54.5%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PYVEKLLPLSL 13
: | | | | | | | | | |
Db 153 PVEKLLIETKL 163

RESULT 7
US-08-540-804-4
Sequence 4, Application US/08540804
Patent No. 5919666
GENERAL INFORMATION:
APPLICANT: Young, Richard A.
APPLICANT: Koleske, Anthony J.
APPLICANT: Thompson, Craig M.
APPLICANT: Chao, David M.
TITLE OF INVENTION: No. 5919666 Factors Which Modify Gene
TITLE OF INVENTION: Transcription and Methods of Use Therefor
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/540,804
FILING DATE: 11-OCT-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/521,872
FILING DATE: 21-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/218,265
FILING DATE: 25-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH194-03A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 687 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-540-804-4

Query Match 52.1%; Score 36.5; DB 2; Length 687;
Best Local Similarity 38.1%; Pred. No. 2.2e+02;
Matches 8; Conservative 5; Mismatches 1; Indels 7; Gaps 1;

OY 1 SNP-----PVEKLLPLSLK 14
||| |::|::|:
DB 88 SNPHETYSMPLEQLPIILR 108

RESULT 8

US-08-218-265-4
Sequence 4, Application US/08218265
Patent No. 5922585
GENERAL INFORMATION:
APPLICANT: Young, Richard A.
APPLICANT: Koleske, Anthony J.
APPLICANT: Thompson, Craig M.
TITLE OF INVENTION: No. 5922585el Factors Which Modify Gene
TITLE OF INVENTION: Transcription and Methods of Use Thereof
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: US
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/218, 265
FILING DATE: 25-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WHI94-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 687 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-218-265-4

Query Match 52.1%; Score 36.5; DB 2; Length 687;
Best Local Similarity 38.1%; Pred. No. 2.2e+02;
Matches 8; Conservative 5; Mismatches 1; Indels 7; Gaps 1;

OY 1 SNP-----PVEKLLPLSLK 14
||| |::|::|:
DB 88 SNPHETYSMPLEQLPIILR 108

RESULT 9
US-08-521-872-4
Sequence 4, Application US/08521872
Patent No. 6015682

GENERAL INFORMATION:
APPLICANT: Young, Richard A.
APPLICANT: Koleske, Anthony J.
APPLICANT: Thompson, Craig M.
APPLICANT: Chao, David M.
TITLE OF INVENTION: No. 6015682el Factors Which Modify Gene
TITLE OF INVENTION: Transcription and Methods of Use Therefor
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/521, 872
FILING DATE: 31-AUG-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/218,265
FILING DATE: 25-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WHI94-03A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 687 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-521-872-4

Query Match 52.1%; Score 36.5; DB 3; Length 687;
Best Local Similarity 38.1%; Pred. No. 2.2e+02;
Matches 8; Conservative 5; Mismatches 1; Indels 7; Gaps 1;

OY 1 SNP-----PVEKLLPLSLK 14
||| |::|::|:
DB 88 SNPHETYSMPLEQLPIILR 108

RESULT 10
US-08-590-399-4
Sequence 4, Application US/08590399
Patent No. 6214588
GENERAL INFORMATION:
APPLICANT: Young, Richard A.
APPLICANT: Koleske, Anthony J.
APPLICANT: Thompson, Craig M.
APPLICANT: Chao, David M.
TITLE OF INVENTION: No. 6214588el Factors Which Modify Gene
TITLE OF INVENTION: Transcription and Methods of Use Therefor
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/590,399
FILING DATE: 26-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/540,804
FILING DATE: 11-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/521,872
FILING DATE: 31-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/218,265
FILING DATE: 25-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH194-03A3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 687 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-590-399-4

Query Match
Best Local Similarity 52.1%; Score 36.5; DB 4; Length 687;
Matches 8; Conservative 5; Mismatches 1; Indels 7; Gaps 1;

OY 1 SNP-----PVEKILPLSLK 14
DB 88 SNPEHIGSMPLDPLILIR 108

RESULT 11
US-08-936-165A-368
Sequence 368, Application US/08936165A
Patent No. 6348582
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nichols, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: Pol. 6348582el Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmli, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 368:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-936-165A-368

Query Match
Best Local Similarity 51.4%; Score 36; DB 4; Length 104;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SNPEVKILPL 10
DB 17 SNPEILPLIP 26

RESULT 12
US-09-362-318-2
Sequence 2, Application US/09362318
Patent No. 6077993
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Maize Repair Protein Orthologue-1 and
FILE REFERENCE: 0934
CURRENT APPLICATION NUMBER: US/09/362,318
CURRENT FILING DATE: 1999-07-27
EARLIER APPLICATION NUMBER: 60/099,280
EARLIER FILING DATE: 1998-09-03
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 315
TYPE: PRT
ORGANISM: Zea mays
US-09-362-318-2

Query Match
Best Local Similarity 51.4%; Score 36; DB 3; Length 315;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 NPPEVKILPL 10
DB 72 NPPEILPLIP 80

RESULT 13
US-07-688-352C-16
Sequence 16, Application US/07688352C
Patent No. 5527896
GENERAL INFORMATION:
APPLICANT: Wiegler, Michael H.
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: Cloning by Complementation and Related
TITLE OF INVENTION: Processes
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark

STREET: Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/688,352C
FILING DATE: 19910419
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25447
REFERENCE/DOCKET NUMBER: 27805/30197
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 464 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-688-352C-16

Query Match 51.4%; Score 36; DB 1; Length 464;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 NPVEKLLP 10
Db 144 DPPIESLLP 152

RESULT 14
US-08-474-379C-16
Sequence 16, Application US/08474379C
Patent No. 5977305
GENERAL INFORMATION:
APPLICANT: Wigler, Michael H.
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
TITLE OF INVENTION: PROCESSES
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,379C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/206,188

FILING DATE: 01-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/688,352
FILING DATE: 19-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 27866/32771
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 464 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-379C-16

Query Match 51.4%; Score 36; DB 2; Length 464;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 NPVEKLLP 10
Db 144 DPPIESLLP 152

RESULT 15
US-09-146-249A-16
Sequence 16, Application US/09146249A
Patent No. 6069240
GENERAL INFORMATION:
APPLICANT: Wigler, Michael H.
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: Cloning by Complementation and Related
TITLE OF INVENTION: Processes
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/146,249A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 27866/32771
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 464 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-146-249A-16

Query Match 51.4%; Score 36; DB 3; Length 464;

Best Local Similarity 55.6%; Pred. NO. 1.7e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 2 NPPVERKLP 10
:|:|:|:|
Db 144 DPPIESLIP 152

Search completed: January 12, 2003, 04:18:44
Job time : 26 secs

This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2003, 04:15:10 ; Search time 17 Seconds
(without alignments)
15.977 Million cell updates/sec

Title: HANDENTERED
Perfect score: 70
Sequence: 1 snpyekllpislk 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues
Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	55.7	48	10	US-09-764-847-905
2	38	54.3	221	10	US-09-925-299-967
3	38	54.3	282	10	US-09-815-242-5576
4	38	54.3	292	10	US-09-815-242-12440
5	37	52.9	283	10	US-09-815-242-10199
6	37	52.9	467	10	US-09-804-551B-10
7	36	51.4	104	10	US-09-939-980-368
8	36	51.4	188	10	US-09-734-569-134
9	36	51.4	203	10	US-09-810-997-4
10	36	51.4	291	10	US-09-815-242-11818
11	36	51.4	303	10	US-09-864-761-34235
12	36	51.4	324	10	US-09-746-801A-61
13	36	51.4	333	10	US-09-815-242-11871
14	36	51.4	401	10	US-09-815-242-5506
15	36	51.4	401	10	US-09-815-242-12261
16	36	51.4	528	10	US-09-892-985-2
17	36	51.4	551	10	US-09-815-242-10500
18	35	50.0	93	9	US-10-081-817-32
19	35	50.0	93	12	US-10-052-586-244

20	35	50.0	270	10	US-09-911-826A-8	Sequence 8, Appli
21	35	50.0	513	9	US-09-738-626-3645	Sequence 3645, Ap
22	35	50.0	550	10	US-09-934-323-6	Sequence 6, Appli
23	35	50.0	603	9	US-09-764-868-705	Sequence 705, App
24	35	50.0	649	9	US-09-984-245-197	Sequence 197, App
25	35	50.0	783	10	US-09-825-144-2	Sequence 2, Appli
26	35	50.0	829	10	US-09-825-144-4	Sequence 4, Appli
27	35	50.0	835	10	US-09-934-323-2	Sequence 2, Appli
28	35	50.0	836	10	US-09-934-323-5	Sequence 5, Appli
29	35	50.0	1263	10	US-09-971-309-6	Sequence 6, Appli
30	34	48.6	69	10	US-09-864-761-42307	Sequence 42307, A
31	34	48.6	123	10	US-09-764-847-509	Sequence 509, App
32	34	48.6	258	10	US-09-994-485-16	Sequence 16, Appli
33	34	48.6	259	9	US-09-832-292-3	Sequence 3, Appli
34	34	48.6	294	9	US-09-738-626-6177	Sequence 6177, Ap
35	34	48.6	328	10	US-09-864-761-35526	Sequence 35526, A
36	34	48.6	463	9	US-10-155-613-1	Sequence 1, Appli
37	34	48.6	479	10	US-09-881-752A-312	Sequence 312, App
38	34	48.6	538	10	US-09-815-242-11125	Sequence 11125, A
39	34	48.6	640	9	US-10-060-019-30	Sequence 30, Appli
40	34	48.6	1146	9	US-09-832-292-10	Sequence 10, Appli
41	34	48.6	1146	10	US-09-824-734-2	Sequence 2, Appli
42	34	48.6	1146	10	US-09-994-485-6	Sequence 6, Appli
43	33	47.1	57	10	US-09-864-761-43961	Sequence 43961, A
44	33	47.1	142	10	US-09-764-847-975	Sequence 975, App
45	33	47.1	149	10	US-09-864-761-43311	Sequence 43311, A

ALIGNMENTS

RESULT 1
US-09-764-847-905
Sequence 905, Application US/09764847
Patient No. US20020132767A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC009
CURRENT APPLICATION NUMBER: US/09/764,847
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 2003
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 905
LENGTH: 48
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (30)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (44)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-905

Query Match 55.7% ; Score 39; DB 10; Length 48;
Best Local Similarity 80.0% ; Pred. No. 2.7;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PVEKILPLSL 13
Db 33 PLERLLPLSL 42

RESULT 2
US-09-925-299-967
Sequence 967, Application US/09925299
Patient No. US20020055627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

```
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 967
LENGTH: 221
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-299-967
```

```
Query Match 54.3% Score 38; DB 10; Length 221;
Best Local Similarity 61.5%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 NNPVEKLPLSL 13
    | | | | |
Db 7 SRPPLTPLSL 19
```

```
RESULT 3
US-09-815-242-5576
Sequence 5576, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5576
LENGTH: 282
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5576
```

```
Query Match 54.3% Score 38; DB 10; Length 282;
Best Local Similarity 80.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 2 NNPVEKLPL 11
    | | | | |
Db 71 NNPVEKLPL 80
```

```
RESULT 4
US-09-815-242-12440
Sequence 12440, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12440
LENGTH: 292
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12440
```

```
Query Match 54.3% Score 38; DB 10; Length 292;
Best Local Similarity 80.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 2 NNPVEKLPL 11
    | | | | |
Db 77 NNPVEKLPL 86
```

```
RESULT 5
US-09-815-242-10199
Sequence 10199, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
```

```

; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10199
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10199

Query Match
Best Local Similarity 70.0%; Score 37; DB 10; Length 283;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 NPPEVKLPL 11
DB 73 NHPITLPL 82

RESULT 6
US-09-804-551B-10
; Sequence 10, Application US/09804551B
; Patent No. US20020056151A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Receptors for peptides from insects
; FILE REFERENCE: Le A 34 394
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: DE 100 13 618.4
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-804-551B-10

Query Match
Best Local Similarity 52.9%; Score 37; DB 10; Length 467;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SNPEVKLPL 12
DB 388 SNPIPLPL 399

RESULT 7
US-09-939-980-368
; Sequence 368, Application US/09939980
; Patent No. US20020082234A1
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; Burnham, Martin
; Hodgson, John
; Knowles, David
; Lonetto, Michael
; Nicholas, Richard
; Pratt, Julie
; Reichard, Richard
; Rosenberg, Martin
; Ward, Judith
; TITLE OF INVENTION: No. US20020082234A1el Prokaryotic Polynucleotides,
; Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
```

```

; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,980
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/936,165
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmil, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 368:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 368:
US-09-939-980-368

Query Match
Best Local Similarity 51.4%; Score 36; DB 10; Length 104;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNPEVKLPL 10
DB 17 SNPIPLPL 26

RESULT 8
US-09-734-569-134
; Sequence 134, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reinold, Andreas
; APPLICANT: Clippus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; APPLICANT: Reski, Ralf
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involv
; FILE REFERENCE: BASF-NAE-1332-99-US
; CURRENT APPLICATION NUMBER: US/09/734,569
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/171,101
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1
; SEQ ID NO 134
; LENGTH: 188
; TYPE: PRT
```

ORGANISM: Physcomitrella patens
US-09-734-569-134

Query Match
Best Local Similarity 51.4%; Score 36; DB 10; Length 188;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NPPVKLLP 10
| | | | |
Db 9 NPPVKLLP 17

RESULT 9

US-09-810-997-4
; Sequence 4, Application US/09810997
; Patent No. US20020007501A1
; GENERAL INFORMATION:
; APPLICANT: Song, Xiaoling
; APPLICANT: Fan, Hao
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: RECEPTORS FOR HYPERSENSITIVE RESPONSE ELICITORS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 21829/62
; CURRENT APPLICATION NUMBER: US/09/810,997
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/191,649
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: 60/250,710
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Oryza
US-09-810-997-4

Query Match
Best Local Similarity 51.4%; Score 36; DB 10; Length 203;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 PPEVKLLPL 13
| | | | |
Db 78 PPEVKLLPL 88

RESULT 10

US-09-815-242-11818
; Sequence 11818, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlisen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11818
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11818

Query Match
Best Local Similarity 51.4%; Score 36; DB 10; Length 291;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 NPPVKLLPL 11
| | | | |
Db 75 NPPVKLLPL 84

RESULT 11

US-09-864-761-34235
; Sequence 34235, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aegm1a-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117

```

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34235
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL049538.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 9.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 8.7
; OTHER INFORMATION: EXPRESSED IN HB1100, SIGNAL = 7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.4
; OTHER INFORMATION: SWISSPROT HIT: P97680, EVALUATE 2.00e-03
; US-09-864-761-34235

```

```

Query Match
Best Local Similarity 51.4%; Score 36; DB 10; Length 303;
Pred. No. 72;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 NPPEVKLIP 10
:|:|:|:|:|
DB 291 DPPIESLIP 299

```

```

RESULT 12
US-09-746-801A-61
; Sequence 61, Application US/09746801A
; Patent No. US20020083494A1
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
; FILE REFERENCE: 1505-34357
; CURRENT APPLICATION NUMBER: US/09/746,801A
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Brassica sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (188)..(188)
; OTHER INFORMATION: Xaa = uncertain amino acid residue
; NAME/KEY: misc_feature
; LOCATION: (194)..(194)
; OTHER INFORMATION: Xaa = uncertain amino acid residue
; NAME/KEY: misc_feature
; LOCATION: (196)..(196)
; OTHER INFORMATION: Xaa = uncertain amino acid residue
; NAME/KEY: misc_feature
; LOCATION: (214)..(214)
; OTHER INFORMATION: Xaa = uncertain amino acid residue
; US-09-746-801A-61

```

```

Query Match
Best Local Similarity 51.4%; Score 36; DB 10; Length 324;
Pred. No. 78;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 NPPEVKLIP 10
:|:|:|:|:|
DB 237 NPPEVKLIP 245

```

```

RESULT 13
US-09-815-242-11871
; Sequence 11871, Application US/09815242
; Patent No. US20020061569A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11871
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-815-242-11871

```

```

Query Match
Best Local Similarity 51.4%; Score 36; DB 10; Length 333;
Pred. No. 81;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 3 PYVEKLIPLSL 13
:|:|:|:|:|
DB 260 PVEESMAPLWL 270

```

```

RESULT 14
US-09-815-242-5506
; Sequence 5506, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

```

```

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5506
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5506

```

```

Query Match      51.4%; Score 36; DB 10; Length 401;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 SNPPVEKLLP 10
    ||||| 1:1
DB 109 SNPPILPLIP 118

```

```

RESULT 15
US-09-815-242-12261
; Sequence 12261, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes In
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12261
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12261

```

```

Query Match      51.4%; Score 36; DB 10; Length 401;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 SNPPVEKLLP 10
    ||||| 1:1
DB 109 SNPPILPLIP 118

```

Search completed: January 12, 2003, 04:19:15
 Job time : 19 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2003, 04:12:01 ; Search time 40 Seconds
(without alignments)
33.647 Million cell updates/sec

Title: HANDENTERED

Perfect score: 70
Sequence: 1 snnpvexklplslk 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	67.1	889	2 T47311	hypothetical prote
2	42	60.0	327	2 D70346	conserved hypothet
3	42	60.0	333	2 T29208	hypothetical prote
4	42	60.0	366	2 T47360	hypothetical prote
5	42	60.0	386	2 A39732	H+-transporting tw
6	42	60.0	450	2 G96749	hypothetical prote
7	42	60.0	663	2 T26835	hypothetical prote
8	42	60.0	1075	2 B96508	hypothetical prote
9	42	60.0	1350	2 T42697	hypothetical prote
10	41	58.6	376	2 S27976	H+-transporting tw
11	41	58.6	593	2 E81277	hypothetical prote
12	41	58.6	649	2 C81275	hypothetical prote
13	41	58.6	653	2 B81277	hypothetical prote
14	40	57.1	105	2 A12020	hypothetical prote
15	40	57.1	288	2 S18438	hypothetical prote
16	40	57.1	1115	2 T11614	sporulation protei
17	39	53.7	316	1 PWTCG	probable poly(A)-s
18	39	53.7	316	1 R5RT10	H+-transporting tw
19	39	55.7	316	1 R5RT10	acidic ribosomal p
20	39	55.7	317	1 R5HUP0	acidic ribosomal p
21	39	55.7	317	1 R5M810	acidic ribosomal p
22	39	55.7	317	1 R5PEP0	acidic ribosomal p
23	39	55.7	475	2 AE3036	slrcheme synthase
24	39	55.7	488	2 F98249	slrcheme synthase
25	39	55.7	653	2 A46362	amyloid precursor-
26	39	55.7	799	2 A82703	phosphoenolpyruvat
27	38	54.3	186	2 C82029	probable adhesin c
28	38	54.3	217	2 C86350	protein F8K.12 [I
29	38	54.3	284	2 F84590	probable heat shoc

30	38	54.3	292	2 A89823	conserved hypothet
31	38	54.3	312	2 T21351	hypothetical prote
32	38	54.3	315	2 E75382	hypothetical prote
33	38	54.3	358	1 PWRMG	H+-transporting tw
34	38	54.3	371	2 T13021	hypothetical prote
35	38	54.3	420	2 A84864	hypothetical prote
36	38	54.3	577	2 C84936	DNA primase [impor
37	38	54.3	738	2 F96701	hypothetical prote
38	38	54.3	1131	2 S22266	FUN30 protein - ye
39	38	54.3	2670	2 T37919	GCN1 homolog - fis
40	38	54.3	2672	2 A48126	translation activa
41	37	52.9	158	2 B90842	hypothetical prote
42	37	52.9	166	2 T27540	hypothetical prote
43	37	52.9	181	2 T14538	probable ethylene
44	37	52.9	227	2 AB2008	hypothetical prote
45	37	52.9	283	2 A99967	hypothetical prote

ALIGNMENTS

RESULT 1
T47311
hypothetical protein T32A11.100 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47311
R:Reger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24459
A:Accession: T47311
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-889 <RIE>
A:Cross-references: EMBL:AL138653
A:Experimental source: cultivar Columbia; BAC clone T32A11
C:Genetics:
A:Map position: 3
A:Introns: 88/3; 282/3; 455/3; 497/3; 547/3; 569/3; 607/3; 629/3; 667/3; 694/3; 716/3
A>Note: T32A11.100

Query Match 67.1%; Score 47; DB 2; Length 889;
Best Local Similarity 75.0%; Pred. No. 4.8;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNNPVEKPLPLS 12
DB 552 SNPSVEKVLPLN 563

RESULT 2
D70346
conserved hypothetical protein aq_512 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: D70346
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.

Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; NCID:98196666; PMID:9537320
A:Accession: D70346
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-327 <NOF>
A:Cross-references: GB:AE000693; NID:g2983148; PIDN:AAC06763.1; PID:g2983161; GB:AE00
A:Experimental source: strain VFS
C:Genetics:
A:Gene: aq_512

Query Match 60.0%; Score 42; DB 2; Length 327;
Best Local Similarity 63.6%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SNPEKILLPL 11
 DB 311 SNPEKILLPL 321

RESULT 3

T29208
 hypothetical protein T20F5.4 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T29208

R:Du, Z.; Le, T.T.
 submitted to the EMBL Data Library, November 1996
 A:Description: The sequence of *C. elegans* cosmid T20F5.
 A:Reference number: Z20588

A:Accession: T29208

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-333 <DN>

A:Cross-references: EMBL:U00442; PIDN:AA37665.1; GSPDB:GN00019; CESP:T20F5.4

C:Genetics:

A:Gene: CESP:T20F5.4

A:Map position: 1

A:Introns: 35/3; 103/2; 177/2; 191/3; 240/3; 314/2

Query Match

Best Local Similarity 60.0%; Score 42; DB 2; Length 333;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 NPPEKILLPLSLK 14

DB 272 SPVRRALPIELK 284

RESULT 4

T47360
 hypothetical protein F7M19.30 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T47360

R:Nyakatura, G.; Farcman, B.; Dauner, D.; Steer, W.; Holland, R.; Weichselgartner, M.;

Mayer, K.F.X.

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24458

A:Accession: T47360

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-366 <NTA>

A:Cross-references: EMBL:AL138643

A:Experimental source: cultivar Columbia; BAC clone F7M19

C:Genetics:

A:Map position: 3

A:Introns: 98/3; 140/3; 189/3; 211/3; 245/3; 272/3; 299/3; 317/3

Query Match

Best Local Similarity 60.0%; Score 42; DB 2; Length 366;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SNPEKILLPLS 12

DB 194 SNQSVKLLPLN 205

RESULT 5

A39732
 H+-transporting two-sector ATPase (EC 3.6.3.14) gamma-2 chain precursor, chloroplast - *A*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 03-Jun-2002

C:Accession: A39732

R:Inohara, N.; Iwamoto, A.; Moriyama, Y.; Shimomura, S.; Maeda, M.; Futai, M.

J. Biol. Chem. 266, 7333-7338, 1991

A:Title: Two genes, atpC1 and atpC2, for the gamma subunit of *Arabidopsis thaliana* ch

A:Reference number: A39732; MUID:91210234; PMID:1826905

A:Accession: A39732

A:Molecule type: DNA

A:Residues: 1-386 <INO>

A:Cross-references: GB:M61741

A>Note: the authors translated the codon GAA for residue 192 as Asp, GAT for residue

C:Genetics:

A:Gene: atpC2

A:Introns: #status absent

C:Superfamily: H+-transporting ATP synthase gamma chain

C:Keywords: ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; th

F:1-60/Domain: transmembrane (chloroplast) #status predicted <TNP>

F:61-386/Product: H+-transporting ATP synthase gamma chain 2 #status predicted <MAT>

F:260-266/Dissulfide bonds: #status predicted

Query Match

Best Local Similarity 60.0%; Score 42; DB 2; Length 386;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 SNPEKILLPLSLK 14

DB 243 SDPVITHTLPLSMK 256

RESULT 6

G96749
 hypothetical protein F28P22.27 [Imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002

C:Accession: G96749

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G96749

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-450 <STO>

A:Cross-references: GB:AE005173; MID:96648174; PIDN:AAF21174.1; GSPDB:GN00141

C:Genetics:

A:Gene: F28P22.27

A:Map position: 1

C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match

Best Local Similarity 60.0%; Score 42; DB 2; Length 450;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SNPEKILLPLS 12

DB 26 TNPPEKILLPLS 37

RESULT 7

T26835
 hypothetical protein Y43FA.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26835

R:Matthews, L.

submitted to the EMBL Data Library, September 1997

A:Reference number: Z20775

A:Accession: T26835

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-663 <MIL>
A:Cross-references: EMBL:Z99271; NID:e1062263; PIDN:CAB16471.1; GSPDB:GN00021; CESP:Y43F
A:Experimental source: clone Y43FA4
C:Genetics:
A:Gene: CESP:Y43FA4.1
A:Map position: 3
A:Introns: 29/3; 120/2; 159/3; 187/2; 273/2; 310/3; 360/3; 393/2; 423/2; 446/2; 567/3

Query Match
Best Local Similarity 60.0%; Score 42; DB 2; Length 663;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 SNPEVKLPLSLK 14
:::|||||:::
Db 60 TSPVEKFAFLRIQ 73

RESULT 8
B96508
hypothetical protein T12C22.16 [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: B96508
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MIMD:21016719; PMID:1130712
A:Accession: B96508
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1075 <STO>
A:Cross-references: GB:AE005173; NID:g8655999; PIDN:AAF78272.1; GSPDB:GN00141
C:Genetics:
A:Gene: T12C22.16
A:Map position: 1

Query Match
Best Local Similarity 60.0%; Score 42; DB 2; Length 1075;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SNPEVKLPLSLK 12
:::|||||:::
Db 569 SNOSEVKLPLSLN 580

RESULT 9
T42697
hypothetical protein DKFZp434N0335.1 - human (fragment)
N:Alternate names: hypothetical protein DKFZp434D0815.1; hypothetical protein DKFZp434I01
C:Species: Homo sapiens (man)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000
C:Accession: T42697; T42651; T42672; T46321
R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, November 1999
A:Reference number: Z22232
A:Accession: T42697
A:Molecule type: mRNA
A:Residues: 1-1350 <POD>
A:Cross-references: EMBL:AL133040
A:Experimental source: adult testis; clone DKFZp434N0335
R:Boeckler, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, November 1999
A:Reference number: Z22230
A:Accession: T42651
A:Molecule type: mRNA

A:Residues: 323-1350 <BLO1>
A:Cross-references: EMBL:AL133048
A:Experimental source: adult testis; clone DKFZp434D0815
C:Accession: T42672
A:Molecule type: mRNA
A:Residues: 435-1350 <BLO2>
A:Cross-references: EMBL:AL133072
A:Experimental source: adult testis; clone DKFZp434I0617
R:Duesterhoef, A.; Lauber, J.; Mewes, H.W.; Well, B.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23036
A:Accession: T46321
A:Molecule type: mRNA
A:Residues: 725-1350 <DUE>
A:Cross-references: EMBL:AL137622
A:Experimental source: adult testis; clone DKFZp434I2412
C:Genetics:
A:Note: DKFZp434N0335.1; DKFZp434D0815.1; DKFZp434I0617.1; DKFZp434I2412.1
C:Superfamily: human retinoblastoma binding protein 2

Query Match
Best Local Similarity 60.0%; Score 42; DB 2; Length 1350;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPEVKLPL 11
:::|||||:::
Db 1033 PPEVKLPL 1041

RESULT 10
S27976
H+-transporting two-sector ATPase (EC 3.6.3.14) gamma chain precursor, chloroplast -
C:Species: Pisum sativum (garden pea)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 03-Jun-2002
C:Accession: S27976; S18951
R:Napier, J.A.; Hoeglund, A.S.; Plant, A.L.; Gray, J.C.
Plant Mol. Biol. 20, 737-741, 1992
A:Title: Chloroplast import of the precursor of the gamma subunit of pea chloroplast
A:Reference number: S27976; MIMD:93081734; PMID:1450388
A:Accession: S27976
A:Molecule type: mRNA
A:Residues: 1-376 <NAP>
A:Cross-references: EMBL:X63604; NID:g20653; PIDN:CAA45150.1; PID:g20654
C:Genetics:
A:Gene: atpC
C:Superfamily: H+-transporting ATP synthase gamma chain
C:Keywords: ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; th
F:1-52/Domain: transit peptide (chloroplast) #status predicted <TNP>
F:53-376/Product: H+-transporting ATP synthase gamma chain #status predicted <MAT>
F:250-256/Disulfide bonds: #status predicted

Query Match
Best Local Similarity 58.6%; Score 41; DB 2; Length 376;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 SNPEVKLPLSLK 14
:::|||||:::
Db 233 SNPIIHTLPLSPK 246

RESULT 11
E81277
hypothetical protein Cj1336 [Imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: E81277
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chl
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Bar
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; MIMD:20150912; PMID:10688204
A:Accession: E81277
A>Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-593 <PAR>
 A:Cross-references: GB:AL139078; GB:AL111168; NID:96968723; PIDN:CAB73763.1; PID:9696877
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: Cj1336
 C:Superfamily: Campylobacter jejuni hypothetical protein Cj1337

Query Match 58.6%; Score 41; DB 2; Length 593;
 Best Local Similarity 75.0%; Pred. No. 36;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Oy 2 NPPEVKLPLSL 13
 Db 427 NKPEKLEPLSL 438

RESULT 12
 C81275
 hypothetical protein Cj1318 [Imported] - Campylobacter jejuni (strain NCTC 11168)
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C:Accession: C81275

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrett
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf
 A:Reference number: AB1250; MUID:20150912; PMID:10688204
 A:Accession: C81275
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-649 <PAR>
 A:Cross-references: GB:AL139078; GB:AL111168; NID:96968723; PIDN:CAB73745.1; PID:9696875
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: Cj1318
 C:Superfamily: Campylobacter jejuni hypothetical protein Cj1337

Query Match 58.6%; Score 41; DB 2; Length 649;
 Best Local Similarity 75.0%; Pred. No. 39;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Oy 2 NPPEVKLPLSL 13
 Db 483 NKPEKLEPLSL 494

RESULT 13
 B81277
 hypothetical protein Cj1333 [Imported] - Campylobacter jejuni (strain NCTC 11168)
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C:Accession: B81277

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrett
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf
 A:Reference number: AB1250; MUID:20150912; PMID:10688204
 A:Accession: B81277
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-653 <PAR>
 A:Cross-references: GB:AL139078; GB:AL111168; NID:96968723; PIDN:CAB73760.1; PID:9696876
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: Cj1333
 C:Superfamily: Campylobacter jejuni hypothetical protein Cj1337

Query Match 58.6%; Score 41; DB 2; Length 653;
 Best Local Similarity 75.0%; Pred. No. 40;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Oy 2 NPPEVKLPLSL 13

Db 483 NKPEKLEPLSL 494

RESULT 14
 AI2020
 hypothetical protein all1719 [Imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AI2020

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AI2020
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-105 <KUN>
 A:Cross-references: GB:BA000019; PIDN:BA873418.1; PID:917130809; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all1719

Query Match 57.1%; Score 40; DB 2; Length 105;
 Best Local Similarity 58.3%; Pred. No. 77;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Oy 3 PPEVKLPLSLK 14
 Db 39 PPEKLEPLQIK 50

RESULT 15
 S18438
 sporulation protein spoIVFB - Bacillus subtilis
 N:Alternate names: pro-sigma-K processing proteinase spoIVFB
 C:Species: Bacillus subtilis
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
 C:Accession: S18438; C69714

R:Cutting, S.; Roels, S.; Losick, R.
 J. Mol. Biol. 221, 1237-1256, 1991
 A:Title: Sporulation operon spoIVF and the characterization of mutations that uncoupl
 A:Reference number: S18437; MUID:92046062; PMID:1942049
 A:Accession: S18438
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-288 <CUN>

A:Cross-references: EMBL:X59528; NID:940170; PIDN:CAA42107.1; PID:9580934
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C.; Bron, S.; Brouillet, S.; Krogth, S.; Kumano, M.; Kuriita, K.; Lapidus, A.; Lardino
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabelt, C.; Ferrari,
 Nature 380, 249-256, 1997
 A:Authors: Foulger, D.R.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
 lech, J.; Hartwood, C.R.; Henuat, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kuriita, K.; Lapidus, A.; Lardino
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
 Y. M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl
 A:Authors: Schleich, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Se
 deuchli, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchida
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
 A:Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
 A:Reference number: A69580; MUID:96044033; PMID:9384377
 A:Accession: C69714
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-288 <KUN>
 A:Cross-references: GB:Z99118; GB:AL009126; NID:92635200; PIDN:CAB14757.1; PID:el1840
 A:Experimental source: strain 168
 C:Genetics:

A:Gene: SPOIVR
A:Start codon: TTG
C:Keywords: Transmembrane protein

Query Match 57.1%; Score 40; DB 2; Length 288;
Best Local Similarity 61.5%; Pred. No. 24;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 NPYEKLLPLSLK 14
| :|||||:
Db 218 NRELEKLLPLTVK 230

Search completed: January 12, 2003, 04:18:11
Job time : 48 secs

This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2003, 01:58:03 ; Search time 21 Seconds

(without alignments)
27.651 Million cell updates/sec

Title: HANDENTERED
Perfect score: 70
Sequence: 1 snppveklplslk 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	60.0	386	1	ATP2_ARATH
2	41	58.6	376	1	ATPC_PRA
3	40	57.1	288	1	SP4G_BACSU
4	40	57.1	1115	1	YAA4_SCHPO
5	39	55.7	117	1	SP21_PAEPO
6	39	55.7	309	1	RLAO_BOVIN
7	39	55.7	316	1	ATPG_SYNP6
8	39	55.7	316	1	RLAO_CHICK
9	39	55.7	317	1	RLAO_DROME
10	39	55.7	317	1	RLAO_HUMAN
11	39	55.7	317	1	RLAO_ICPU
12	39	55.7	317	1	RLAO_MOUSE
13	39	55.7	317	1	RLAO_RAT
14	39	55.7	653	1	APPL_MOUSE
15	38	54.3	311	1	RLAO_CAEEL
16	38	54.3	316	1	RLAO_PLAF8
17	38	54.3	358	1	ATPG_CHLRE
18	38	54.3	577	1	PRIM_BUCAI
19	38	54.3	1131	1	YAB9_YEAST
20	38	54.3	2670	1	YAO5_SCHPO
21	38	54.3	2672	1	GCN1_YEAST
22	37	52.9	219	1	SSL_DROME
23	37	52.9	282	1	YEDU_ECOLI
24	37	52.9	364	1	ATPG_SPROL
25	37	52.9	393	1	HVC3_HETER
26	37	52.9	394	1	LPXB_SYNY3
27	37	52.9	438	1	HVC5_HETER
28	37	52.9	461	1	HVC6_HETER
29	37	52.9	497	1	IRFS_MOUSE
30	37	52.9	1165	1	PEX6_PICPA
31	36.5	52.1	687	1	SRB4_YEAST
32	36.5	52.1	1181	1	NKX1_RAT
33	36	51.4	226	1	PRH1_HUMAN

34	36	51.4	283	1	YK26_YEAST	P36139	saccharomyc
35	36	51.4	301	1	NHAR_ECOLI	P10087	eschlerichia
36	36	51.4	314	1	THP1_SFKA	P16472	shope fibro
37	36	51.4	319	1	RLAO_BRARE	O9P9V0	brachydanio
38	36	51.4	373	1	ATP1_ARATH	O01908	arabidopsis
39	36	51.4	377	1	ATPG_TOBAC	P29790	nicotiana t
40	36	51.4	529	1	ACH2_HUMAN	O15822	homo sapien
41	36	51.4	572	1	CY42_TRYBB	O99396	trypanosoma
42	36	51.4	1042	1	EF3_PNECA	P29551	pneumocysti
43	35	50.0	93	1	UGRI_HUMAN	O96P11	homo sapien
44	35	50.0	110	1	RPOZ_MYCLE	O9CG66	mycobacteri
45	35	50.0	110	1	RPOZ_MYCTU	P71660	mycobacteri

ALIGNMENTS

RESULT 1	ATP2_ARATH	STANDARD:	PRT:	386 AA.
AC	O01909;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DE	ATP synthase gamma chain 2, chloroplast precursor (EC 3.6.3.14).			
GN	ATPC2 OR ATG15700 OR F7H2.4.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91210234; PubMed=1826905;			
RA	Itohara N., Iwamoto A., Moriyama Y., Shimomura S., Maeda M.,			
RA	Futai M.,			
RT	"Two genes, atpC1 and atpC2, for the gamma subunit of Arabidopsis			
RT	thaliana chloroplast ATP synthase.";			
RL	J. Biol. Chem. 266:7333-7338(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV, Columbia;			
RX	MEDLINE=21016719; PubMed=11130712;			
RA	Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,			
RA	White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,			
RA	Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,			
RA	Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,			
RA	Dunn P., Elgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,			
RA	Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,			
RA	Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,			
RA	Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,			
RA	Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,			
RA	Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marzilli A.,			
RA	Miltscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,			
RA	Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,			
RA	Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,			
RA	Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,			
RA	Uttterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,			
RA	Wu D., Yu G., Fraser C.M., Venter U.C., Davis R.W.,			
RT	"Sequence and analysis of chromosome 1 of the plant Arabidopsis			
RT	thaliana.";			
RL	Nature 408:816-820(2000).			
CC	- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON			
CC	GRADIENT ACROSS THE MEMBRANE. THE GAMMA CHAIN IS BELIEVED TO BE			
CC	IMPORTANT IN REGULATING ATPASE ACTIVITY AND THE FLOW OF PROTONS			
CC	THROUGH THE CF(0) COMPLEX.			
CC	- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +			
CC	H(+)(Out).			
CC	- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC			
CC	CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE			
CC	SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)			
CC	HAS THREE MAIN SUBUNITS: A, B AND C.			
CC	- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.			

```

CC -1- INDUCTION: BY LIGHT.
CC -1- SIMILARITY: BELONGS TO THE ATPASE GAMMA CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
DR EMBL: M61742; AAA32833.1; -
DR EMBL: AC034256; AAF82140.1; -
DR PIR: A39732; A39732.
DR InterPro: IPR000131; ATPase_gamma.
DR Pfam: PF00231; ATP-synt. 1.
DR PRINTS: PR00126; ATPASEGAMMA.
DR TIGRFS: TIGR01146; ATPsyn_F1gamma; 1.
DR PROSITE: PS00153; ATPASE_GAMMA; 1.
KM ATP synthetase; Chloroplast; Thylakoid; Membrane; CF(1);
KM Hydroxylase; Hydrogen ion transport; Transl. peptide; Multigene family.
FT TRANSIT 1 60 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 1 60 ATP SYNTHASE GAMMA CHAIN 2.
FT ACT_SITE 61 386 BY SIMILARITY.
FT DISULFID 149 149 BY SIMILARITY.
FT SEQUENCE 260 266 BY SIMILARITY.
SQ SEQUENCE 386 AA; 42679 MW; 2E0CE2F47EBAC3A CRC64;

Query Match
Best Local Similarity 60.0%; Score 42; DB 1; Length 386;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 SNPPVKLPLSLK 14
DB 243 SDPVHTLPLSLK 256
ID 1:1:1111:1
AC P28552;
ATPq_PEA STANDARD: PRT; 376 AA.
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP synthase gamma chain, chloroplast precursor (EC 3.6.3.14).
GN ATPC.
OS Pisum sativum (Garden pea).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
CC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OK NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93081734; PubMed=1450388;
RA Napier J.A., Hglund A.S., Plant A.L., Gray J.C.;
RT "Chloroplast import of the precursor of the gamma subunit of pea
RT chloroplast ATP synthase."
RL Plant Mol. Biol. 20:737-741(1992).
CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC GRADIENT ACROSS THE MEMBRANE. THE GAMMA CHAIN IS BELIEVED TO BE
CC IMPORTANT IN REGULATING ATPASE ACTIVITY AND THE FLOW OF PROTONS
CC THROUGH THE CF(0) COMPLEX.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(in) = ADP + phosphate +
CC H(+)(out).
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -1- SIMILARITY: BELONGS TO THE ATPASE GAMMA CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
DR EMBL: X63604; CAA45150.1; -
DR PIR: S18951; S18951.
DR PIR: S27976; S27976.
DR InterPro: IPR000131; ATPase_gamma.
DR InterPro: IPR01230; Ptenyl_site.
DR Pfam: PF00231; ATP-synt. 1.
DR PRINTS: PR00126; ATPASEGAMMA.
DR TIGRFS: TIGR01146; ATPsyn_F1gamma; 1.
DR PROSITE: PS00153; ATPASE_GAMMA; FALSE NEG.
KM ATP synthetase; Chloroplast; Thylakoid; Membrane; CF(1);
KM Hydroxylase; Hydrogen ion transport; Transl. peptide.
FT TRANSIT 1 52 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 53 376 ATP SYNTHASE GAMMA CHAIN.
FT ACT_SITE 133 133 BY SIMILARITY.
FT DISULFID 250 256 BY SIMILARITY.
SQ SEQUENCE 376 AA; 41410 MW; 6976834B12F3CD18 CRC64;

Query Match
Best Local Similarity 58.6%; Score 41; DB 1; Length 376;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 SNPPVKLPLSLK 14
DB 233 SNPIHTLPLSLK 246
ID 1:1:1111:1
AC P26937;
SP4G_BACSU STANDARD: PRT; 288 AA.
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Stage IV sporulation protein PB (EC 3.4.24.-).
GN SPOIVFB OR BOFB.
OS Bacillus subtilis.
CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OK NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=92046062; PubMed=1942049;
RA Cutting S.M., Roels S., Losick R.;
RT "Sporulation operon spoIVF and the characterization of mutations that
RT uncouple mother-cell from forespore gene expression in Bacillus
RT subtilis."
RL J. Mol. Biol. 221:1237-1256(1991).
RN [2]
RP MUTAGENESIS OF HIS-43; GLU-44 AND HIS-47.
RX MEDLINE=20270168; PubMed=10809718;
RA Yu Y.-T., Kroos L.;
RT "Evidence that SPOIVFB is a novel type of membrane metalloprotease
RT governing intercompartmental communication during Bacillus subtilis
RT sporulation."
RL J. Bacteriol. 182:3305-3309(2000).
CC -1- FUNCTION: IMPLICATED IN THE COUPLING OF MOTHER CELL TO FORESPORE
CC GENE EXPRESSION. REQUIRED FOR SPORE FORMATION. PROBABLE
CC PROTEOLYTIC THAT PROCESSES THE PRO-SIGMA K FACTOR.
CC -1- SUBUNIT: FORMS AN HETERODIMER WITH SPOIVFA (PROBABLY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
CC surrounding the forespore protoplast (probable).
CC NOT REQUIRED UNTIL THE STAGE IV, OF SPOULATION.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M0B.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X59528; CAA42107.1; -;
DR EMBL: Z99118; CAB14757.1; -;
DR PIR: S18438; S18438.
DR MEROPS: M50.002; -;
DR Subtilist: BG10332; spoIVFB.
KW Hydrolyase; Metalloprotease; Zinc; Sporulation; Transmembrane;
KW Complete proteome.

FT DOMAIN 1 10 MOTHER-CELL CYTOPLASM (PROBABLE).
FT TRANSMEM 11 30 POTENTIAL.
FT DOMAIN 31 31 FORESPORE INTERMEMBRANE SPACE (PROBABLE).
FT TRANSMEM 32 56 POTENTIAL.
FT DOMAIN 57 83 MOTHER-CELL CYTOPLASM (PROBABLE).
FT TRANSMEM 84 105 POTENTIAL.
FT TRANSMEM 106 128 FORESPORE INTERMEMBRANE SPACE (PROBABLE).
FT TRANSMEM 127 146 POTENTIAL.
FT DOMAIN 147 161 MOTHER-CELL CYTOPLASM (PROBABLE).
FT TRANSMEM 162 178 POTENTIAL.
FT DOMAIN 179 179 FORESPORE INTERMEMBRANE SPACE (PROBABLE).
FT TRANSMEM 180 199 POTENTIAL.
FT DOMAIN 200 288 MOTHER-CELL CYTOPLASM (PROBABLE).
FT METAL 43 43 ZINC (CATALYTIC) (PROBABLE).
FT ACT_SITE 44 44 PROBABLE.
FT METAL 47 47 ZINC (CATALYTIC) (PROBABLE).
FT METAL 43 43 ZINC (CATALYTIC) (PROBABLE).
FT MUTAGEN 44 44 H->F: LOSS OF ACTIVITY.
FT MUTAGEN 44 44 E->A/O: LOSS OF ACTIVITY.
FT MUTAGEN 44 44 E->D: PARTIAL ACTIVITY.
FT MUTAGEN 47 47 H->F: LOSS OF ACTIVITY.
SQ SEQUENCE 288 AA; 33640 MW; 81DB758DC215C916 CRC64;

Query Match 57.1%; Score 40; DB 1; Length 288;
Best Local Similarity 61.5%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 NPVEKLLPLSLK 14
| : ||||| : :
Db 218 NRELEKLLPLTVK 230

RESULT 4
YAA4_SCHPO STANDARD; PRT; 1115 AA.
ID YAA4_SCHPO 009798;
AC 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C2267.04 in chromosome I.
GN SPAC2267.04.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Spouris J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbittowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Mamut R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt K., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cervetti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Sinapovskii G.V., Ussery D., Barrell B.G., Nurse P.;
RA "The genome sequence of *Schizosaccharomyces pombe*."
RL Natural 415:871-880(2002).
CC -1- SIMILARITY: TO YEAST PAN2 AND TO C. ELEGANS F31E3.4.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Z54328; CAA91128.1; -;
DR InterPro: IPR000520; Exonuclease.
DR InterPro: IPR001394; UCH-2.
DR Pfam: PF00443; UCH-2; 1.
DR Pfam: PF00929; Exonuclease; 1.
DR SMART: SM00479; EXOIII; 1.
DR SMART: PS50235; UCH_2_3; 1.
DR PROSITE: PS50235; UCH_2_3; 1.
KW Hypothetical protein.
SQ SEQUENCE 1115 AA; 126919 MW; EEF6F15E1897269D CRC64;

Query Match 57.1%; Score 40; DB 1; Length 1115;
Best Local Similarity 54.5%; Pred. No. 44;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 NPVEKLLPLSLK 12
| : ||| : :
Db 866 NPINKIQLPLT 876

RESULT 5
SP21_PAEPO STANDARD; PRT; 117 AA.
ID SP21_PAEPO 032720;
AC 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Anti-sigma F factor antagonist (Stage II sporulation protein AA).
GN SPOI1A.
OS Paenibacillus polymyxa (Bacillus polymyxa).
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
OX NCBI_TaxID=1406;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 36;
RX MEDLINE=97409954; PubMed=9266669;
RA Park S.G., Yuckin M.D.;
RT "Sequencing and phylogenetic analysis of the *spoI1A* operon from
RT diverse *Bacillus* and *Paenibacillus* species."
RL Gene 194:25-33(1997).
CC -1- FUNCTION: IN THE PHOSPHORYLATED FORM IT COULD ACT AS AN ANTI-ANTI-
CC SIGMA FACTOR THAT COUNTERACTS SPOI1B AND THUS RELEASES SIGMA F
CC FROM INHIBITION (BY SIMILARITY).
CC -1- PFM: PHOSPHORYLATED BY SPOI1B ON A SERINE RESIDUE (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ANTI-SIGMA-FACTOR ANTAGONIST FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL: L47358; AAB81184.1; -
 DR HSP: P10727; IAD2.
 DR InterPro: IPR003658; Antisig_antgst.
 DR InterPro: IPR002645; STAS.
 DR Pfam: PF01740; STAS; 1.
 DR Prodom: PD005210; Antisig_antgst; 1.
 DR TIGRFAMs: TIGR00377; ant_ant_sig; 1.
 DR PROSITE: PS50801; STAS; 1.
 DR Sporulation: Phosphorylation.
 KW DOMAIN 3 STAS.
 FT MOD_RES 58 PHOSPHORYLATION (BY SIMILARITY).
 SO SEQUENCE 117 AA; 13341 MW; 7D66D2AABD788B0A CRC64;
 Query Match 55.7%; Score 39; DB 1; Length 117;
 Best Local Similarity 63.6%; Pred. No. 6;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 OY 2 NPPEVKLPLS 12
 DB 84 NPPVRLDMS 94
 RESULT 6
 ID RLA0_BOVIN STANDARD; PRT; 309 AA.
 AC 095140; 018788; Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE 60S acidic ribosomal protein P0 (L10E) (Fragment).
 GN RPLP0.
 OS Bos taurus (bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE OF 1-50 FROM N.A.
 RC TISSUE=Adrenal cortex;
 RA Mandirola S.J., Pepper M.S.;
 RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
 RP [2]
 RP SEQUENCE OF 8-309 FROM N.A.
 RC TISSUE=Adorta;
 RA Lillestiek B., Nocha M., Umanaky V., Benner A., Lin Y., Ziegler R.,
 RA Navroth P.P., Schlirmacher V.;
 RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
 CC -1 FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
 CC OF E COLI PROTEIN L10
 CC -1 SUBUNIT: P0 FORMS A PENTAMERIC COMPLEX BY INTERACTION WITH
 CC DIMERS OF P1 AND P2.
 CC -1 SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL: U73275; AAB17671.1; -
 DR EMBL: AF013214; AAB65436.1; -
 DR InterPro: IPR001813; 60S_ribosomal.
 DR InterPro: IPR001790; Ribosomal_L10.
 DR Pfam: PF00428; 60S_ribosomal; 1.
 DR Pfam: PF00466; Ribosomal_L10; 1.

KW Ribosomal protein; Phosphorylation.
 FT NON_TER 1
 SO SEQUENCE 309 AA; 33257 MW; C8A70665EEF7EB6C CRC64;
 Query Match 55.7%; Score 39; DB 1; Length 309;
 Best Local Similarity 70.0%; Pred. No. 17;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 1 SNPEVKLPL 10
 DB 62 NNPVRLDMS 71
 RESULT 7
 ID ATPG_SYN6 STANDARD; PRT; 316 AA.
 AC P08450; Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ATP synthase gamma chain (EC 3.6.3.14).
 GN ATPG OR ATPC.
 OS Synechococcus sp. (strain PCC 6301) (Anacystis nidulans).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 OX NCBI_TaxID=1139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87311713; PubMed=3041005;
 RX Cozens A.L., Walker J.E.;
 RT "The organization and sequence of the genes for ATP synthase subunits
 RT in the cyanobacterium Synechococcus 6301. Support for an
 RT endosymbiotic origin of chloroplasts.";
 RL J. Mol. Biol. 194:359-383(1987).
 RN [2]
 RP SEQUENCE OF 306-316 FROM N.A.
 RX MEDLINE=88326273; PubMed=2843173;
 RA Cozens A.L., Walker J.E.;
 RT "Expression of a gene encoding a novel ferredoxin in the
 RT cyanobacterium Synechococcus 6301.";
 RL Biochem. J. 252:563-569(1988).
 CC -1 FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
 CC GRADIENT ACROSS THE MEMBRANE. THE GAMMA CHAIN IS BELIEVED TO BE
 CC IMPORTANT IN REGULATING ATPASE ACTIVITY AND THE FLOW OF PROTONS
 CC THROUGH THE CF(0) COMPLEX.
 CC -1 CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
 CC H(+)(Out).
 CC -1 SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
 CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
 CC HAS THREE MAIN SUBUNITS: A, B AND C.
 CC -1 SIMILARITY: BELONGS TO THE ATPASE GAMMA CHAIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL: X03502; CAA28929.1; -
 DR PIR: S10832; PWYCG.
 DR InterPro: IPR000131; ATPase_gamma.
 DR Pfam: PF00231; ATP-synt; 1.
 DR PRINTS: PR00126; ATPASGAMMA.
 DR TIGRFAMs: TIGR01146; ATPSyn_Flgamma; 1.
 DR PROSITE: PS00153; ATPASE_GAMMA; 1.
 KW ATP synthetase; CF(1); Hydrogen ion transport; Hydrolyase.
 SO SEQUENCE 316 AA; 34806 MW; 7451EB6D23DADB0B CRC64;
 Query Match 55.7%; Score 39; DB 1; Length 316;
 Best Local Similarity 72.7%; Pred. No. 17;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

OY      1 SNPEVKLPL 11
      1111111111
Db      183 SNPEVKLPL 193

RESULT 8
RLAO_CHICK
ID      RLAO_CHICK      STANDARD;      PRT;      316 AA.
AC      P47826;
DT      01-FEB-1996 (Rel. 33, Last sequence update)
DT      01-FEB-1996 (Rel. 33, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DE      60S acidic ribosomal protein P0 (L10E).
GN      RPLP0.
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Gallus.
OX      NCBI_TaxID=9031;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-White leghorn; TISSUE-Liver;
RX      MEDLINE=96057559; Pubmed=7549959;
RA      Wang H., Meury L., Plasmoneault S., Morais R.;
RT      "Chicken acidic ribosomal phosphoprotein P0: Isolation and molecular
RT      characterization of cDNA clones.";
RL      Biochem. Mol. Biol. Int. 36:595-604(1995).
CC      -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC      OF E.COLI PROTEIN L10.
CC      -1- SUBUNIT: P0 FORMS A PENTAMERIC COMPLEX BY INTERACTION WITH
CC      DIMERS OF P1 AND P2.
CC      -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L28704; AAC38020.1; -.
DR      InterPro: IPR001813; 60S_rribosomal.
DR      InterPro: IPR001790; Ribosomal_L10.
DR      Pfam: PF00428; 60S_rribosomal; 1.
DR      Pfam: PF00466; Ribosomal_L10; 1.
KM      Ribosomal protein; Phosphorylation.
SQ      SEQUENCE 316 AA; 34285 MW; AF0DB9/DE87584B5 CRC64;

Query Match      55.7%; Score 39; DB 1; Length 316;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY      1 SNPEVKLPL 10
      1111111111
Db      71 NNPALEKLP 80

RESULT 9
RLAO_DROME
ID      RLAO_DROME      STANDARD;      PRT;      317 AA.
AC      P19889; O9VNV9;
DT      01-FEB-1991 (Rel. 17, Created)
DT      01-FEB-1991 (Rel. 17, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      60S acidic ribosomal protein P0 (Deoxyribonuclease (Apyriminic or
DE      apyrimidinase)) (EC 4.2.99.18) (Apyrimidin-apyrimidinic endonuclease).
GN      RPP0 OR APE OR AP3 OR CG7490.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC      Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC      Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

```

```

OX      NCBI_TaxID=7227;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=89261760; Pubmed=2471063;
RX      Kelley M.R., Vennuppal S., Harless J., Deutsch W.A.;
RT      "Antibody to a human DNA repair protein allows for cloning of a
RT      Drosophila cDNA that encodes an apurinic endonuclease.";
RL      Mol. Cell. Biol. 9:965-973(1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN-Berkeley;
RX      MEDLINE=20196006; Pubmed=10731132;
RX      Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA      Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA      Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA      April J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA      Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA      Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA      Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA      Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA      de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA      Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA      Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA      Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA      Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA      Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA      Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA      Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA      Kiskkel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA      Lasse P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA      Liu X., Mattei B., McIntosh T.C., McLeod M.P., Mephorson D.,
RA      Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA      Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA      Nelson D.R., Nelson K.A., Nixon K.A., Nuskern D.R., Pacleb J.M.,
RA      Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA      Rebert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA      Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA      Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA      Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA      Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA      Williams S.M., Wodade T., Worley K.C., Wu D., Yang S., Yao O.A.,
RA      Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA      Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA      Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT      "The genome sequence of Drosophila melanogaster.";
RL      Science 287:2185-2195(2000).
RN      [3]
RP      SIMILARITY TO RIBOSOMAL PROTEIN P0.
RX      MEDLINE=9136151; Pubmed=1870984;
RX      Grabowski D.T., Deutsch W.A., Derda D., Kelley M.R.;
RT      "Drosophila AP3, a presumptive DNA repair protein, is homologous to
RT      human ribosomal associated protein P0.";
RL      Nucleic Acids Res. 19:4297-4297(1991).
RN      [4]
RP      DNA REPAIR ACTIVITY.
RX      MEDLINE=97086697; Pubmed=8932386;
RX      Yacoub A., Kelley M.R., Deutsch W.A.;
RT      "Drosophila ribosomal protein P0 contains apurinic/aprimidinic
RT      endonuclease activity.";
RL      Nucleic Acids Res. 24:4298-4303(1996).
CC      -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC      OF E.COLI PROTEIN L10.
CC      -1- CATALYTIC ACTIVITY: The C-O-P bond 3' to the apurinic or
CC      apyrimidinic site in DNA is broken by a beta-elimination reaction,
CC      leaving a 3'-terminal unsaturated sugar and a product with a
CC      terminal 5'-phosphate.
CC      -1- SUBUNIT: P0 FORMS A PENTAMERIC COMPLEX BY INTERACTION WITH
CC      DIMERS OF P1 AND P2 (BY SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.

```

```

CC -1- DEVELOPMENTAL STAGE: ALL STAGES OF DEVELOPMENT. A LARGER
CC TRANSCRIPT IS RESTRICTED TO THE EMBRYONIC AND EARLY LARVAL
CC STAGES.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M25772; AAA53372.1; -
DR EMBL: AE003596; AAF51807.1; -
DR PIR: A30223; R5FEP0.
DR FLYBASE: FBgn0000100; RFP0.
DR InterPro: IPR001813; 60s_ribosomal.
DR InterPro: IPR001790; Ribosomal_L10.
DR Pfam: PF00428; 60s_ribosomal; 1.
DR Pfam: PF00466; Ribosomal_L10; 1.
DR Ribosomal protein; Phosphorylation; Hydrolyase; Lyase; Nuclease;
DR Endonuclease; DNA repair; Nuclear protein.
DR MOD_RES 304 363DDSDDCFELEIF CRC64;
SQ SEQUENCE 317 AA; 34202 MW; 363DDSDDCFELEIF CRC64;

Query Match 55.7%; Score 39; DB 1; Length 317;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNRPVEKLIP 10
Db 71 NNPALEKLIP 80

RESULT 10
RLAO_HUMAN STANDARD; PRT; 317 AA.
AC P05388; Q9BVK4;
DR 01-NOV-1988 (Rel. 09, Created)
DR 01-NOV-1988 (Rel. 09, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60s acidic ribosomal protein P0 (L10E).
GN RPLP0.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86122131; Pubmed=3323886;
RA Rich B.E.; Steitz J.A.;
RT "Human acidic ribosomal phosphoproteins P0, P1, and P2: analysis of
RT cDNA clones, in vitro synthesis, and assembly.";
RL Mol. Cell. Biol. 7:4065-4074(1987).
RN [2]
RP SEQUENCE FROM N.A.
RA Bradshaw H.; Ozersky P.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA TISSUE=Brain, Cervix, Colon, Lung, Lymph, Muscle, and Pancreas;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 218-310 FROM N.A.
RX MEDLINE=98248690; Pubmed=9582194;
RA Kameuchi N.; Kawaguchi T.; Rozen S.; Davis E.; Goodman N.;
RA Hudson T.J.; Tanaka T.; Page D.C.;
RT "A map of 75 human ribosomal protein genes.";
RL Genome Res. 8:509-523(1998).
CC -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC OF E. COLI PROTEIN L10.

```

```

CC -1- SUBUNIT: P0 FORMS A PENTAMERIC COMPLEX BY INTERACTION WITH
CC DIMERS OF P1 AND P2.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17885; AAA36470.1; -
DR EMBL: AC004263; AAC05176.1; -
DR EMBL: BC000087; AAH00087.1; -
DR EMBL: BC000345; AAH00345.1; -
DR EMBL: BC000752; AAH00752.1; -
DR EMBL: BC001127; AAH01127.1; -
DR EMBL: BC001834; AAH01834.1; -
DR EMBL: BC003655; AAH03655.1; -
DR EMBL: BC005863; AAH05863.1; -
DR EMBL: BC008092; AAH08092.1; -
DR EMBL: BC008594; AAH08594.1; -
DR EMBL: BC009867; AAH09867.1; -
DR EMBL: BC015173; AAH15173.1; -
DR EMBL: BC015690; AAH15690.1; -
DR EMBL: AB007187; BAA25845.1; -
DR PIR: A27125; R5HUP0.
DR GeneW; HGNC:10371; RPLP0.
DR MIM; 180510; -
DR InterPro: IPR001813; 60s_ribosomal.
DR InterPro: IPR001790; Ribosomal_L10.
DR Pfam: PF00428; 60s_ribosomal; 1.
DR Pfam: PF00466; Ribosomal_L10; 1.
DR Ribosomal protein; Phosphorylation.
FT CONFLICT 246 246 K -> E (IN REF. 3; AAH01127).
SQ SEQUENCE 317 AA; 34273 MW; 255AD25571CS1199 CRC64;

Query Match 55.7%; Score 39; DB 1; Length 317;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNRPVEKLIP 10
Db 71 NNPALEKLIP 80

RESULT 11
RLAO_ICTPU STANDARD; PRT; 317 AA.
AC Q90YX1;
DR 15-JUN-2002 (Rel. 41, Created)
DR 15-JUN-2002 (Rel. 41, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60s acidic ribosomal protein P0 (L10E).
GN RPLP0.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Ictaluridae; Ictalurus.
OX NCBI_Taxid=7998;
RN [1]
RP SEQUENCE FROM N.A.
RA Patterson A.P.; Karsi A.; Liu Z.J.;
RT "Translational machinery of channel catfish: II. Complementary DNA and
RT expression of the complete set of 47 60s ribosomal proteins.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC OF E. COLI PROTEIN L10.
CC -1- SUBUNIT: P0 FORMS A PENTAMERIC COMPLEX BY INTERACTION WITH
CC DIMERS OF P1 AND P2 (By similarity).
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF401551; AAK95123.1; -
DR InterPro: IPR001813; 60s_ribosomal.
DR InterPro: IPR001790; Ribosomal_L10.
DR Pfam: PF00428; 60s_ribosomal_1.
DR Pfam: PF00466; Ribosomal_L10; 1.
DR Ribosomal protein; Phosphorylation.
KW SEQUENCE 317 AA; 34902 MW; 13E2FA0C08EE82AD CRC64;

Query Match 55.7%; Score 39; DB 1; Length 317;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SNPPVERKLP 10
Db 71 NNPALEKILP 80
:|:|||||

RESULT 12
ID RLA0_MOUSE STANDARD; PRT; 317 AA.
AC P14869;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S acidic ribosomal protein P0 (L10E).
GN RPLP0 OR ARBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8936686; PubMed=2771657;
RA Krowczynska A.M., Coutts M., Makrides S., Braverman G.;
RT "The mouse homologue of the human acidic ribosomal phosphoprotein P0:
RT a highly conserved polypeptide that is under translational control.";
RL Nucleic Acids Res. 17:6408-6408(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=21083660; PubMed=11217851;
RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staehli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barst G.,
RA Blake J., Boffelli D., Bojunga M., Carlini P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto T.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Togo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlski S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC OF E.COLI PROTEIN L10.
CC -1- SUBUNIT: P0 FORMS A PENTAMERIC COMPLEX BY INTERACTION WITH
CC DIMERS OF P1 AND P2.

CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X15096; CAA33199.1; -
DR EMBL; AK010267; BAB26807.1; -
DR EMBL; AK012606; BAB28352.1; -
DR PIR: S05305; R5MS10.
DR MGD: MGI:88066; Ardp.
DR InterPro: IPR001813; 60s_ribosomal.
DR InterPro: IPR001790; Ribosomal_L10.
DR Pfam: PF00428; 60s_ribosomal_1.
DR Pfam: PF00466; Ribosomal_L10; 1.
DR Ribosomal protein; Phosphorylation.
KW SEQUENCE 317 AA; 34216 MW; 7985E1D7B235EAD0 CRC64;

Query Match 55.7%; Score 39; DB 1; Length 317;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SNPPVERKLP 10
Db 71 NNPALEKILP 80
:|:|||||

RESULT 13
ID RLA0_RAT STANDARD; PRT; 317 AA.
AC P19945;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 60S acidic ribosomal protein P0 (L10E).
GN RPLP0.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=92075759; PubMed=1742361;
RA Wool I.G., Chan Y.-L., Glueck A., Suzuki K.;
RT "The primary structure of rat ribosomal proteins P0, P1, and P2 and a
RT proposal for a uniform nomenclature for mammalian and yeast ribosomal
RL proteins.";
RL Biochimie 73:861-870(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Myeloid;
RX Loedemel O., Molven A., Houge G.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC OF E.COLI PROTEIN L10.
CC -1- SUBUNIT: P0 FORMS A PENTAMERIC COMPLEX BY INTERACTION WITH
CC DIMERS OF P1 AND P2.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X15096; CAA33199.1; -

```

DR EMBL: Z29530; CAAB2647.1; -.
DR PIR: S08021; RSRT10.
DR PIR: A48397; A48397.
DR PIR: S41396; S41396.
DR InterPro: IPR001813; 60s_ribosomal.
DR InterPro: IPR001790; Ribosomal_L10.
DR Pfam: PF00428; 60s_ribosomal_1.
DR Pfam: PF00466; Ribosomal_L10; 1.
KW Ribosomal protein; Phosphorylation.
FT CONFLICT 281 V -> L (IN REF. 1).
FT CONFLICT 287 A -> L (IN REF. 1).
FT CONFLICT 294 MISSING (IN REF. 1).
SO SEQUENCE 317 AA; 34215 MW; 7B842971C0E027EC CRC64;

Query Match
Best Local Similarity 55.7%; Score 39; DB 1; Length 317;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNPPVKLLP 10
DB 71 NNPALEKLLP 80

RESULT 14
ID APP1_MOUSE STANDARD; PRT; 653 AA.
AC 003157;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Aplyoid-like protein 1 precursor (APLP).
GN APLP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93066322; PubMed=1279693;
RA Masco W., Bupp K., Magendantz M., Gussella J.F., Tanzi R.E.,
RA Solomon F.;
RT "Identification of a mouse brain cDNA that encodes a protein related
RT to the Alzheimer disease-associated amyloid beta protein precursor.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10758-10762(1992).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MAY BE PROCESSED
CC IN THE GOLGI COMPLEX; HOWEVER, IT IS NOT YET CLEAR WHETHER APLP
CC IS SECRETED.
CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L04538; AAA37247.1; -.
DR PIR: A46362; A46362.
DR HSSP: P05067; 1MPV.
DR MGP: MGI:88046; APLP1.
DR InterPro: IPR001868; A4_APP.
DR Pfam: PF02177; A4_EXTRA; 1.
DR PRINTS: PR00203; AMYLIDPA.
DR SMART: SM00006; A4_EXTRA; 1.
DR PROSITE: PS00319; A4_EXTRA; 1.
DR PROSITE: PS00320; A4_INTRA; 1.
DR Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 20
FT CHAIN 21 653
FT DOMAIN 21 583
FT TRANSMEM 584 606
FT POTENTIAL.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.

```

```

FT DOMAIN 607 653
FT DOMAIN 643 646
FT DOMAIN 263 271
FT CARBOHYD 464 464
FT CARBOHYD 554 554
SO SEQUENCE 653 AA; 72751 MW; 56516DC3EA40E4B0 CRC64;

Query Match
Best Local Similarity 55.7%; Score 39; DB 1; Length 653;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PYVEKLLPLSL 13
DB 17 PPLPLPLPLSL 27

RESULT 15
ID RLAO_CAEEL STANDARD; PRT; 311 AA.
AC 093572;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S acidic ribosomal protein P0.
GN RPA-0 OR F25H2.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodietinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilkinson J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-15.
RC STRAIN=Bristol N2;
RX Biol. L., Heid H., Liberatori S., Geller G., Pallini V., Zwilling R.;
RT "Two-dimensional gel electrophoresis of Caenorhabditis elegans
RT homogenates and identification of protein spots by microsequencing.";
RL Electrophoresis 18:557-562(1997).
CC -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC OF E. COLI PROTEIN L10.
CC -1- SUBUNIT: P0 FORMS A PENTAMERIC COMPLEX BY INTERACTION WITH
CC DIMERS OF P1 AND P2.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z79754; CAB02098.1; -.
DR Sienna-2DPAGE: 093572; -.
DR WormPep: F25H2.10; CE09655.
DR InterPro: IPR001813; 60s_ribosomal.
DR InterPro: IPR001790; Ribosomal_L10.
DR Pfam: PF00428; 60s_ribosomal_1.
DR Pfam: PF00466; Ribosomal_L10; 1.
KW Ribosomal protein; Phosphorylation.
FT INTRAMET 0
FT SEQUENCE 311 AA; 33642 MW; 2FA9A35CD24DE0F4 CRC64;

Query Match
Best Local Similarity 54.3%; Score 38; DB 1; Length 311;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNPPVKLLP 10
DB 11:|||||

```

Mon Jan 13 09:24:32 2003

handentered.rsp

Page 9

Db 71 NPSLEKLLP 79

Search completed: January 12, 2003, 04:15:39
Job time : 27 secs

This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2003, 02:06:50 ; Search time 80 Seconds

(without alignments)
36.058 Million cell updates/sec

Title: HANDENTERED

Perfect score: 1 snppvek11p1s1k 14

Sequence: BLOSUP62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

671580

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	100.0	360	4	09P2Y6
2	47	67.1	889	10	09M1M3
3	43	61.4	213	17	08T1D7
4	42	60.0	209	2	09K31
5	42	60.0	209	2	09LBD4
6	42	60.0	327	15	066800
7	42	60.0	333	5	P91479
8	42	60.0	366	10	09M259
9	42	60.0	425	16	08R7J8
10	42	60.0	431	4	094800
11	42	60.0	437	11	08VCO4
12	42	60.0	450	10	09CAH1
13	42	60.0	626	4	09NS27
14	42	60.0	663	5	062446
15	42	60.0	916	4	09U1W7
16	42	60.0	916	4	09U1W7

17	42	60.0	1028	4	09URC7	09ufc7 homo sapien
18	42	60.0	1075	10	09LPE2	09lpe2 arabidopsis
19	42	60.0	1350	4	09UPD3	09upd3 homo sapien
20	42	60.0	1544	4	09Y305	09y305 homo sapien
21	42	60.0	1580	4	095811	095811 homo sapien
22	42	60.0	1681	4	09UGL1	09ugl1 homo sapien
23	42	58.6	374	12	09DH06	09dh06 yaba-like d
24	41	58.6	593	16	09PMW3	09pmw3 campylobact
25	41	58.6	649	16	09PMW1	09pmw1 campylobact
26	41	58.6	653	16	09PMW6	09pmw6 campylobact
27	40	57.1	105	16	08W96	08w96 anabaena sp
28	40	57.1	110	2	09Z3F9	09z3f9 synechococ
29	40	57.1	320	12	08V3M1	08v3m1 swinepox v1
30	40	57.1	349	10	09FTT4	09ftt4 oryza sativ
31	39	55.7	315	5	08WQJ2	08wqj2 spodoptera
32	39	55.7	317	4	09BVK4	09bv44 homo sapien
33	39	55.7	317	11	099154	099154 mus musculu
34	39	55.7	317	13	090YX1	090yx1 ictalurus p
35	39	55.7	343	13	08UW98	08uw98 xenopus lae
36	39	55.7	347	5	09GZC7	09gzc7 trypanosoma
37	39	55.7	366	12	09OBC3	09obc3 yaba monkey
38	39	55.7	367	16	092N15	092n15 rhizobium m
39	39	55.7	442	4	09NUP1	09nup1 homo sapien
40	39	55.7	442	4	08W061	08w061 homo sapien
41	39	55.7	488	16	08U931	08u931 agrobacteri
42	39	55.7	542	12	091FG3	091fg3 chilo iride
43	39	55.7	654	11	08VC38	08vc38 mus musculu
44	39	55.7	717	5	09U422	09u422 drosophila
45	39	55.7	799	16	09PDX0	09pdx0 xylella fas

ALIGNMENTS

RESULT 1
ID 09P2Y6 PRELIMINARY; PRT; 360 AA.
AC 09P2Y6; (TREMUREL. 15, Created)
DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)
DT 01-OCT-2000 (TREMUREL. 21, Last annotation update)
DE Carboxypeptidase B-like protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsunoto A.;
RT "Isolation, molecular cloning, and partial characterization of a novel
carboxypeptidase B from human plasma."
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92042093; PubMed=1939207;
RA Eaton D.L., Malloy B.E., Tsai S., Henzel W., Drayna D.;
RT "Isolation, molecular cloning, and partial characterization of a novel
carboxypeptidase B from human plasma."
DR J. Biol. Chem. 266:21833-21838(1991).
DR EMBL; AB011969; BAA90475.1; -;
DR HSSP; P00730; 2CTC.
DR MEROPS; M14.009; -;
DR InterPro; IPR003146; Propep_M14.
DR InterPro; IPR000834; Zn_carboxypept.
DR Pfam; PF00244; Propep_M14.1.
DR Pfam; PF00246; Zn_carboxypept; 1.
DR PRINTS; PR00765; CRBOXYPTASEA.
SQ SEQUENCE 360 AA; 40935 MW; BF670B2F7437C1CB CRC64;

Query Match 100.0%; Score 70; DB 4; Length 360;
Best Local Similarity 100.0%; Pred. NO. 0.00022;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

*4-identical
c-term. identical*

QY 1 SNPPVEKLPLSLK 14
 DB 347 SNPPVEKLPLSLK 360

RESULT 2

Q9MIM3 PRELIMINARY; PRT; 889 AA.
 AC Q9MIM3;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE Hypothetical 100.0 KDa protein.
 GN T32A11_100.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;

Q9MIM3 PRELIMINARY; PRT; 889 AA.
 AC Q9MIM3;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE Hypothetical 100.0 KDa protein.
 GN T32A11_100.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;

QY 1 SNPPVEKLPLSLK 12
 DB 552 SNPPVEKLPLSLK 563

RESULT 3

Q8RLD7 PRELIMINARY; PRT; 213 AA.
 AC Q8RLD7;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DE Predicted protein.
 GN MA3099.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 NCBI_TaxID=2214;

RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity.";
 RT Genome Res. 12:532-542(2002).
 DR EMBL; AF011013; AAM06472.1; -.
 KW Complete proteome.
 SQ SEQUENCE 213 AA; 24042 MW; 8C7A47C1A28451A CRC64;

Query Match 61.4%; Score 43; DB 17; Length 213;
 Best Local Similarity 50.0%; Pred. No. 9.3;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 PVEKLPLSLK 14
 DB 189 PVEKLPLSLK 200

RESULT 4

Q9KW31 PRELIMINARY; PRT; 209 AA.
 AC Q9KW31;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE HypB4.
 GN HRPB4.
 OS Xanthomonas oryzae pv. oryzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xanthomonas.
 NCBI_TaxID=64187;

Q9KW31 PRELIMINARY; PRT; 209 AA.
 AC Q9KW31;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE HypB4.
 GN HRPB4.
 OS Xanthomonas oryzae pv. oryzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xanthomonas.
 NCBI_TaxID=64187;

QY 1 SNPPVEKLPLSLK 13
 DB 162 SNPPVEKLPLSLK 174

RESULT 5

Q9LBD4 PRELIMINARY; PRT; 209 AA.
 AC Q9LBD4;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE HypB4.
 GN HRPB4.
 OS Xanthomonas oryzae pv. oryzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xanthomonas.
 NCBI_TaxID=64187;

Query Match 60.0%; Score 42; DB 2; Length 209;
 Best Local Similarity 76.9%; Pred. No. 14;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SNPEKILPLSL 13
 ||| | ||| |||
 DB 162 SNPVAKLALSL 174

RESULT 6
 056783
 ID Q56783 PRELIMINARY; PRT; 209 AA.
 AC Q56783;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HRPB4.
 GN HRPB4.
 OS Xanthomonas campestris (pv. vesicatoria).
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xanthomonas.
 OX NCBI_TaxID=341;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=75-3;
 RX MEDLINE=96112739; PubMed=8664494;
 RA Fenselau S., Bonas U.;
 RT "Sequence and expression analysis of the hrp pathogenicity operon of Xanthomonas campestris pv. vesicatoria which encodes eight proteins with similarity to components of the Hrp, Ysc, Spa, and Fli secretion systems."
 RL MOL. Plant Microbe Interact. 8:845-854(1995).
 DR EMBL; U33548; AAB08459.1; -;
 SQ SEQUENCE 209 AA; 22358 MW; CC227C752FB9291 CRC64;

Query Match 60.0%; Score 42; DB 2; Length 209;
 Best Local Similarity 76.9%; Pred. No. 14;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SNPEKILPLSL 13
 ||| | ||| |||
 DB 162 SNPVAKLALSL 174

RESULT 7
 066800
 ID Q66800 PRELIMINARY; PRT; 327 AA.
 AC Q66800;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein A0_512.
 GN A0_512.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., Graham D.E., Overbeek R., Snead M.A., Keller M., Huber R., Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex aeolicus."
 RL Nature 392:353-358(1998).
 DR EMBL; AE000693; AAC06763.1; -;
 DR InterPro; IPR001485; PG/PMW_mutase.
 DR Pfam; PF02878; PGM_PMM.1;
 DR PROSITE; PS00710; PGM_PMM.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 327 AA; 37884 MW; A2200A1CDD7076C7 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 327;
 Best Local Similarity 63.6%; Pred. No. 21;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SNPEKILPLSL 11
 |||| : ||
 DB 311 SNPLPKMPL 321

RESULT 8
 P91479
 ID P91479 PRELIMINARY; PRT; 333 AA.
 AC P91479;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical 37.8 kDa protein.
 GN T20F5.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium."
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Du Z., Le T.T.;
 RT "The sequence of C. elegans cosmid T20F5."
 RL submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission."
 RL submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U80442; AAB37665.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 333 AA; 37825 MW; 8CF145617A9DC8CD CRC64;

Query Match 60.0%; Score 42; DB 5; Length 333;
 Best Local Similarity 53.8%; Pred. No. 22;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 NPVEKILPLSLK 14
 : ||| : || : ||
 DB 272 SPVRRALPLSLK 284

RESULT 9
 09M259
 ID Q9M259 PRELIMINARY; PRT; 366 AA.
 AC Q9M259;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE Hypothetical 40.6 kDa protein.
 GN F7M19.30.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nyakatura G., Partmann B., Dauner D., Steir W., Holland R., Weichselgartner M., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.,

```

RL Nat. Biotechnol. 16:1338-1342(1998).
DR EMBL: AB015348; BA34803.1; -.
DR InterPro: IPR001965; znf_PHD.
DR Pfam: PF00628; PHD; 1.
DR SMART; SM00249; PHD; 1.
FT NON_TER 1
SQ SEQUENCE 431 AA; 48637 MW; 747B327207643996 CRC64;

Query Match
Best Local Similarity 60.0%; Score 42; DB 4; Length 431;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0

Qy 3 PVEKILPL 11
||:||||
Db 147 PLEKILPL 155

RESULT 12
Q8VQC4 Q8VQC4 PRELIMINARY; PRT; 437 AA.
AC Q8VQC4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 49.2 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY TUMOR;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC019446; AAH19446.1; -.
DR InterPro: IPR001965; znf_PHD.
DR Pfam: PF00628; PHD; 2.
DR SMART; SM00249; PHD; 2.
DR Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 437 AA; 49224 MW; F8F0EA766D817E87 CRC64;

Query Match
Best Local Similarity 60.0%; Score 42; DB 11; Length 437;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PVEKILPL 11
||:||||
Db 120 PLEKILPL 128

RESULT 13
Q9CAH1 Q9CAH1 PRELIMINARY; PRT; 450 AA.
AC Q9CAH1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 50.9 kDa protein.
GN F28P22.27.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Euphorbia; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; Pubmed=11130712;
RA Theologis A., Becker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altali H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.

```

RA Dunn P., Etgu P., Feldblum T.V., Peng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kvan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Miliescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utechtack T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).
 DR EMBL; AC010926; AAC51840.1; -;
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR PRODOM; PD000001; Euk_pkinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyTKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Hypothetical protein; Transferase.
 SQ SEQUENCE 450 AA; 50875 MW; 847753292B6AF26 CRC64;

Query Match 60.0%; Score 42; DB 10; Length 450;
 Best Local Similarity 75.0%; Pred. No. 29;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SNPEKILPLS 12
 Db 26 TNPKEKILPLS 37

RESULT 14
 O9NSZ7 PRELIMINARY; PRT; 626 AA.
 AC O9NSZ7;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical 70.4 kDa protein (Fragment).
 GN DPF2P434I2412.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-TESTIS;
 RA Duesterhoeft A., Lauber J., Mewes H.W., Weil B., Wiemann S.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL137622; CAB70847.1; -;
 DR InterPro: IPR001965; Znf_PHD.
 DR Pfam; PF00628; PHD; 2.
 DR SMART; SM00249; PHD; 2.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 626 AA; 70422 MW; 3FE05038C914E22C CRC64;

Query Match 60.0%; Score 42; DB 4; Length 626;
 Best Local Similarity 77.8%; Pred. No. 41;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPEKILPL 11
 Db 309 PPEKILPL 317

RESULT 15
 O62446 PRELIMINARY; PRT; 663 AA.
 AC O62446;
 ID O62446;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Y43FA.1 protein.
 GN Y43FA.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matthews L.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99069613; PubMed-9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z99271; CAB16471.1; -;
 DR MEROPS; M08.002; -;
 DR InterPro: IPR001577; Peptidase_M8.
 DR InterPro: IPR000130; Zn_MTpeptidse.
 DR Pfam; PF01457; Peptidase_M8; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 663 AA; 76050 MW; 814C317228B5ACBA CRC64;

Query Match 60.0%; Score 42; DB 5; Length 663;
 Best Local Similarity 50.0%; Pred. No. 43;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 SNPEKILPLSLK 14
 Db 60 TSPPEKFAPLRIQ 73

Search completed: January 12, 2003, 04:17:18
 Job time : 90 secs

This Page Blank (uspto)